

Detecting the birth and death of finite-time coherent sets

Gary Froyland¹ and Péter Koltai²

¹School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia.

²Institute of Mathematics, Freie Universität Berlin, 14195 Berlin, Germany.

Abstract

Finite-time coherent sets (FTCSs) are distinguished regions of phase space that resist mixing with the surrounding space for some finite period of time; physical manifestations include eddies and vortices in the ocean and atmosphere, respectively. The boundaries of finite-time coherent sets are examples of Lagrangian coherent structures (LCSs). The selection of the time duration over which FTCS and LCS computations are made in practice is crucial to their success. If this time is longer than the lifetime of coherence of individual objects then existing methods will fail to detect the shorter-lived coherence. It is of clear practical interest to determine the full lifetime of coherent objects, but in complicated practical situations, for example a field of ocean eddies with varying lifetimes, this is impossible with existing approaches. Moreover, determining the timing of emergence and destruction of coherent sets is of significant scientific interest. In this work we introduce new constructions to address these issues. The key components are an inflated dynamic Laplace operator and the concept of semi-material FTCSs. We make strong mathematical connections between the inflated dynamic Laplacian and the standard dynamic Laplacian [Fro15], showing that the latter arises as a limit of the former. The spectrum and eigenfunctions of the inflated dynamic Laplacian directly provide information on the number, lifetimes, and evolution of coherent sets.

1. Introduction

Lagrangian methods have proven to be powerful tools for elucidating the transport properties of non-autonomous and time-dependent dynamical systems. Beginning with early approaches [MMP84, RKLW90, Pie91, PY93, HP98] on identifying distinguished transport barriers, in the last fifteen years there has been focus on coherent behaviour. This includes so-called Lagrangian coherent structures (LCSs), again targeting barriers to transport, and for which there are a large variety of approaches to their definition and identification:

a very small sample of this work is [SLM05, RSPB11, AT12, BM12, HBV13, MB14, AP15]. Finite-time coherent sets (FTCSs) [FSM10, Fro13] are mobile regions in the phase space that resist mixing and provide a skeleton around which more complicated dynamics occurs. Despite the moniker “coherent”, these structures are often ephemeral: they emerge, live for some time, and then decay and die.

Lagrangian methods are by their nature concerned with computations that follow trajectories over some specified time interval of interest, rather than combining information across time at a fixed location in phase space as in so-called Eulerian methods (e.g. using sea-surface height as a method of finding ocean eddies [Fu06, CSSdS07]). By following trajectories, Lagrangian methods thus primarily detect structures that are coherent (according to various criteria) *for the dominant part of the time interval under study*. This reliance of Lagrangian coherent structure theory and numerics on objects being coherent throughout (or throughout a large proportion of) the computed flow duration has remained essentially unchanged since their introduction almost two decades ago.

The question of determining when coherent structures are born and when they die is largely unaddressed in the dynamical systems literature. To quote MacMillan *et al.* [MOR20]: “One major shortcoming of these (LCS) techniques, however, is the lack of an objective procedure for identifying time scales of interest, or an ability to characterise the lives, deaths, or age of coherent structures, especially when relevant flow time scales are larger than the time scales associated with coherence.” Several previous studies have investigated lifetimes in the context of ocean eddies, e.g. Froyland *et al.* [FHR⁺12] first identified a suitable timescale and then carried out a series of FTCS computations on time windows sliding forward in time, Andrade *et al.* [ACKBV20] exhaustively search a discretised two-parameter space (t, T) where t is the initial time and T is the flow duration, using these pairs as variable inputs to many separate LCS computations. El Aouni [EA21] identifies the timespans of local rotational motion during each Lagrangian trajectory and then defines an eddy as those trajectories that are close at the beginning and the end of their respective timespans.

We build a theoretical framework to directly tackle this problem, using the successful spectral approach of the dynamic Laplace operator [Fro15, FK20] as a foundation. We time-expand our spatial domain to create an inflated dynamic Laplace operator and allow “time” to become a diffusion process itself. This enables us to relax the strict requirement that coherent sets or coherent structures be exactly material (i.e., follow flow trajectories), and leads to the notion of semi-material FTCS, which naturally allow coherent regions to appear and vanish over time. Our constructions are interpreted from multiple viewpoints: the spectrum of Laplace–Beltrami operators, the trajectories of stochastic differential equations, and properties of dynamic Riemannian metrics.

1.1. Setting and Background

We consider deterministic and stochastically perturbed time-dependent dynamical systems. We work primarily in continuous time, however the ideas and constructions naturally cover the discrete-time case. Let $v : [0, \tau] \times \mathbb{R}^d$ denote a time-dependent velocity field, over a finite time duration $[0, \tau]$; for simplicity we assume that $v(t, \cdot)$ is divergence free for all $t \in [0, \tau]$. Denote by $M \subset \mathbb{R}^d$ a full-dimensional, connected, compact sub-

manifold of \mathbb{R}^d , representing the phase space at time $t = 0$, by $\phi_t : M \rightarrow \mathbb{R}^d$ the flow map generated by v from time 0 to time t , and set $M_t = \phi_t(M)$, $t \in [0, \tau]$. Denote by $\mathcal{P}_t : L^2(M) \rightarrow L^2(M_t)$ the transfer operator for ϕ_t , using Lebesgue as the reference measure. Finite-time coherent sets described by [Fro13, DJM16, FKS20] are constructed by adding small isotropic diffusion to the phase space dynamics. One creates an operator $\mathcal{P}_{\varepsilon,t} : L^2(M) \rightarrow L^2(M_t)$, which solves the Fokker–Planck equation

$$\partial_t f = -\nabla \cdot (fv) + \frac{\varepsilon^2}{2} \Delta f \quad (1)$$

with homogeneous Neumann boundary conditions; that is, $f(t, \cdot) := \mathcal{P}_{\varepsilon,t} f(0, \cdot)$ is a solution to (1). For $t > 0$, the compact operator $\mathcal{P}_{\varepsilon,t}$ has a singular value 1 with unit multiplicity, and there is a gap to the next singular value.

Coherent sets over the interval $[0, \tau]$ are described by level sets of leading singular vectors of $\mathcal{P}_{\varepsilon,\tau}$; in particular at time 0, one considers level sets of the eigenvectors f of $\mathcal{P}_{\varepsilon,\tau}^* \mathcal{P}_{\varepsilon,\tau}$ and at time t level sets of $\mathcal{P}_{\varepsilon,t} f$, see [Fro13, FPG14]. For small ε , these coherent sets are approximately *material* under the purely advective dynamics of ϕ_t , meaning that if A_t is a coherent set at time t , then $A_t \approx \phi_t(\phi_s^{-1}(A_s))$. The family becomes more material as ε is decreased [FPG14].

In the limit as $\varepsilon \rightarrow 0$, for fixed t , $\mathcal{P}_{\varepsilon,t}^* \mathcal{P}_{\varepsilon,t}$ approaches the identity operator, and one can take a singular limit to obtain a *dynamic Laplace operator* [Fro15, FK20, KS21], denoted Δ^D . In this purely deterministic setting, coherent sets at time $t = 0$ are identified as level sets of dominant eigenfunctions of Δ^D [Fro15]. These level sets are *exactly material* under the flow ϕ_t , and represent material sets that stay most coherent under vanishing diffusion. In this work we relax the strict materiality requirement while maintaining purely advective dynamics on $\phi_t(M)$, $t \in [0, \tau]$. This will enable the identification of coherent sets that appear and disappear within some larger time window $[0, \tau]$.

Let $\phi_t^* e$ denote the pullback of the Euclidean metric e from the manifold $\phi_t(M)$ to the manifold M . In the divergence-free setting considered in this paper, the dynamic Laplacian Δ^D on $L^2(M)$ has the form [Fro15]

$$\Delta^D = \frac{1}{\tau} \int_0^\tau \Delta_{\phi_t^* e} dt = \frac{1}{\tau} \int_0^\tau \Delta_{g_t} dt, \quad (2)$$

where from now on we use the notation $g_t := \phi_t^* e$. This is an average of Laplace–Beltrami operators for the Riemannian manifolds (M, g_t) , $t \in [0, \tau]$.

1.2. Relaxing materiality and a new key object

In order to relax materiality, we time-expand the phase space, giving each manifold (M, g_t) its own t -fibre. Topologically, this time-expanded domain is simply

$$\mathbb{M}_0 := \bigcup_{t \in [0, \tau]} \{t\} \times M = [0, \tau] \times M. \quad (3)$$

We define the “co-evolved” spacetime manifold by

$$\mathbb{M}_1 := \bigcup_{t \in [0, \tau]} \{t\} \times \phi_t(M). \quad (4)$$

In \mathbb{M}_0 , a curve corresponding to the trajectory $\{\phi_t(x) : 0 \leq t \leq \tau\} \subset M$ is simply the line $\{(t, x) : 0 \leq t \leq \tau\}$. The canonical mapping from \mathbb{M}_0 to the trajectory manifold \mathbb{M}_1 , associating initial conditions with trajectories, is

$$\Phi : \mathbb{M}_0 \rightarrow \mathbb{M}_1, \quad (t, x) \mapsto (t, \phi_t(x)). \quad (5)$$

Figure 1 illustrates these constructions in two situations: there is a coherent family of sets $\phi_t(A)$ throughout the whole time interval $[0, \tau]$ (upper row) and a coherent family of sets $\phi_t(A_t)$ that is born at time τ_1 and extinguished at time τ_2 (lower row).

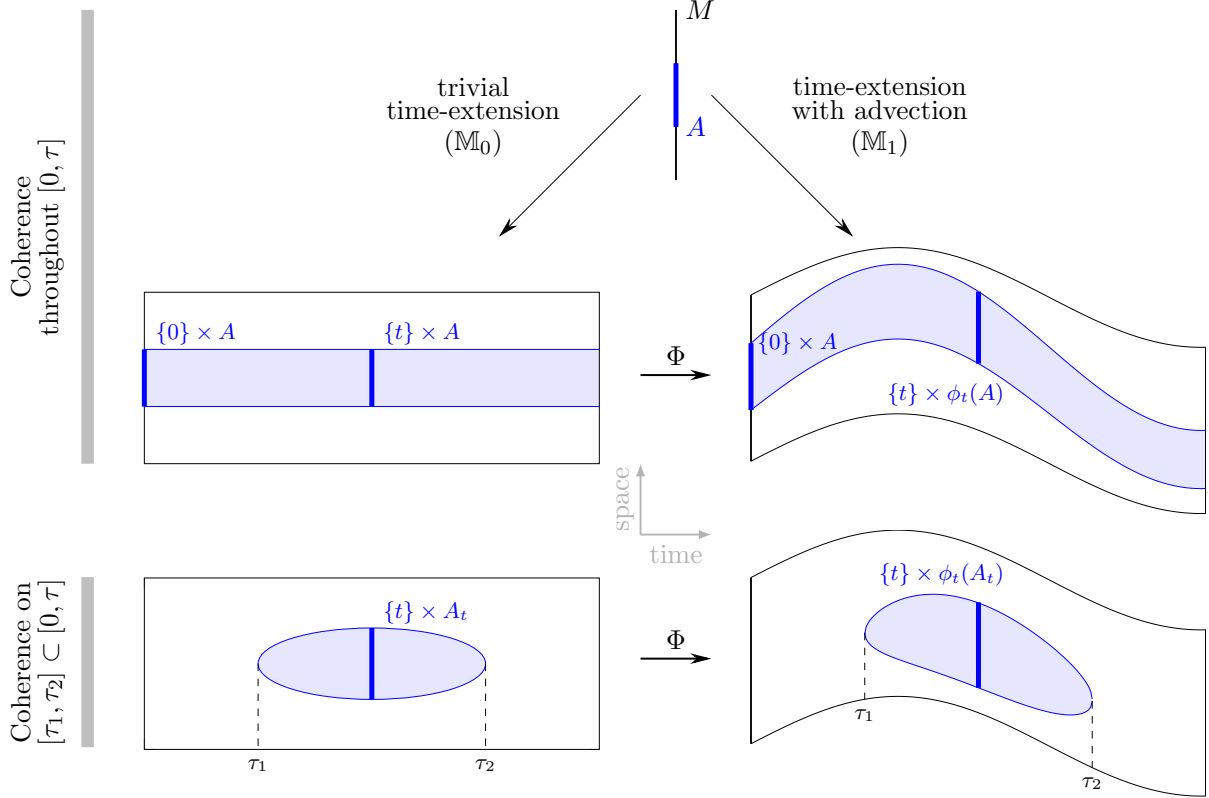


Figure 1.: *Time-expanded constructions and fully vs partially present coherence.* The top of the diagram shows the situation where there is (for simplicity of presentation) a single coherent set $A \subset M$ present throughout the time interval $[0, \tau]$, shown as a dark vertical blue line at the very top of the figure. Upper left: by trivial copying in time we obtain the pale blue horizontal strip $[0, \tau] \times A \subset \mathbb{M}_0$. Upper right: by evolving A forward in time with the dynamics from time 0 to time τ we trace out the pale blue set $\bigcup_{t \in [0, \tau]} \{t\} \times \phi_t(A) \subset \mathbb{M}_1$. The lower row of the diagram concerns the situation where there is a coherent set present *only for part of the time interval*, say a subinterval $[\tau_1, \tau_2] \subset [0, \tau]$. Lower right: following the dynamics, a coherent set appears at τ_1 from a small expanding core, exists for a while, and then shrinks and dissipates completely at τ_2 . Lower left: We pull back the lower right image to time $t = 0$ using the inverse of Φ .

By considering the Euclidean metric on each $\phi_t(M)$, the above constructions naturally suggest a metric on \mathbb{M}_0 . At a point $(t, x) \in \mathbb{M}_0$, we define local distances by the metric

with coordinate representation

$$\begin{pmatrix} 1 & 0 \\ 0 & D\phi_t(x)^\top D\phi_t(x) \end{pmatrix}, \quad (6)$$

where the lower right block is the local matrix representation of $g_t = \phi_t^* e$ at $x \in M$. We denote by G_0 the metric on \mathbb{M}_0 given pointwise by (6).

Our key new object is the *inflated dynamic Laplace operator*, which we briefly now describe, with further details to follow. Consider the Laplace–Beltrami operator $\Delta_{G_0} : L^2(\mathbb{M}_0, G_0) \rightarrow L^2(\mathbb{M}_0, G_0)$; for the moment we delay the discussion of boundary conditions. Because of the time-fibered structure (6) of the metric G_0 , we may write $\Delta_{G_0} F(t, \cdot) = \partial_{tt} F(t, \cdot) + \Delta_{g_t} F(t, \cdot)$. In fact, we will consider a family of Laplace–Beltrami operators with a parameter $a > 0$,

$$\Delta_{G_{0,a}} F(t, \cdot) = a^2 \partial_{tt} F(t, \cdot) + \Delta_{g_t} F(t, \cdot). \quad (7)$$

We will show that these operators interpolate between the dynamic Laplacian Δ^D (in the $a \rightarrow \infty$ limit), whose level sets of eigenfunctions form *exactly material* families of coherent sets, and a purely non-dynamic Laplace–Beltrami operator (when $a = 0$), whose level sets of eigenfunctions need not have any material properties. Thus, the parameter a interpolates the material requirement from strictly material to non-material. By selecting an appropriate a , the eigenfunctions of $\Delta_{G_{0,a}}$ will identify (i) time intervals of strong and weak mixing, and (ii) coherent sets within the time intervals of weak mixing.

We note that time-expansion has been used in the context of transfer operators to find coherent sets of periodic [FK17], finite-time aperiodic [FKS20] and aperiodic [GD20] flows. There are several differences between these works and our current constructions, including, but not limited to: (i) we do not require coherent behaviour throughout the flow duration being considered, (ii) we work with Laplace–Beltrami operators instead of transfer operators, (iii) our analysis is carried out on \mathbb{M}_0 , rather than the co-evolved time-expanded manifold, (iv) we consider time as a diffusion process, instead of it increasing with a constant speed. Moreover, in addition to the stochastic trajectory and transfer operator interpretations in [FK17, FKS20], we also provide a differential-geometric perspective. Other work arising from the dynamic Laplacian includes [KK20, KS21, SKJ21], where the emphasis is on the time-averaged processes generated by the dynamic Laplacian in the initial time slice on M .

Laplace-spectral approaches [GDGGG⁺13] to analysing multilayer networks [DDSRC⁺13, BBC⁺14, KAB⁺14] share some structural similarities to (7), where diffusion occurs both within and across network layers. In the particular case of a two-layer network, [GDGGG⁺13] study the dependence of the spectrum of a “supra-Laplacian” on the coupling strength. This is formally similar to the construction of [FKN⁺19, equation (7)], which considers multiple layers. Our results concerning the behaviour of the spectrum and eigenfunctions of $\Delta_{G_{0,a}}$ with varying diffusion strength a should carry over to multilayer networks to describe the analogous behaviour with varying interlayer coupling strength and connect to graph-based versions of the dynamic Laplace operator [FK15].

Finally, the birth and death of a coherent set represents a structural change in the dynamics. An unrelated type of structural change for almost-invariant and coherent sets

is the crossing of eigenvalues of the transfer operator, which is sometimes, but not always, associated with the bifurcation of the sets. Case studies that consider these types of perturbations include [JMM04] (autonomous), [GRSK12] (periodic), and [BGT20, NPGR21] (non-autonomous). However, we note that bifurcations are not particularly prevalent, for example if a spectral value is currently isolated from other spectral values, [AFJ] shows that the eigenvalues and eigenfunctions of the dynamic Laplacian vary differentially for small perturbations of the flow duration, or of the underlying dynamics.

Outline In section 2.1 we provide independent motivation for the above geometric construction using stochastic trajectories, and then connect this to the geometry on \mathbb{M}_1 and \mathbb{M}_0 in sections 2.2 and 2.3, respectively. In section 3 we derive results interpolating between material and non-material coherence. Section 3.1 briefly recaps the dynamic Laplace operator and section 3.2 shows that the dynamic Laplacian Δ^D arises from the inflated dynamic Laplacian $\Delta_{G_{0,a}}$ in the $a \rightarrow \infty$ limit, by invoking the theory of averaging. The dynamic theory of Cheeger and Sobolev constants for Δ^D is linked to classical notions of these constants on the Riemannian manifold \mathbb{M}_0 in section 3.3. We characterise the behavior of the spectrum of the inflated dynamic Laplacian $\Delta_{G_{0,a}}$ in section 3.4, connecting it with the spectrum of the dynamic Laplacian. In section 4 this theoretical information is synthesised into a practical approach to find coherent sets with lifetimes shorter than the full flow duration. A reduced PDE corresponding to (7) where all spatial information is collapsed is derived in section 5.1, enabling a comparison of instantaneous coherent set decay at time t with average decay across $[0, \tau]$. In idealised coherent and mixing regimes, section 5.2 provides fine detail on the behaviour of the time-fibre norms of the eigenfunctions of the inflated dynamic Laplacian. We develop a trajectory-based numerical scheme based on a specialised finite element method in section 6, and illustrate our theory via an example in section 7.

2. From diffusion to geometry

Recall that deterministic trajectories are represented in \mathbb{M}_0 as straight lines parallel to the time axis, and they can be uniquely parametrized by their initial conditions $(0, x_0) \in \{0\} \times M$. In this section we interpret the paths generated by the SDE associated with the inflated dynamic Laplace operator $\Delta_{G_{0,a}}$ in (7) on \mathbb{M}_0 . These paths are driven by a pure diffusion process and in section 2.3 we will show that these paths *independently* jump *along* and *between* trajectories $\{(t, x_0) \mid t \in [0, \tau]\}$ of the deterministic flow in \mathbb{M}_0 . In a dynamical sense it is natural to first consider the dynamics on its “true”, co-evolved space, so we begin our analysis with a process on \mathbb{M}_1 in sections 2.1 and 2.2, and then pull these constructions back with Φ to \mathbb{M}_0 , in section 2.3.

2.1. Trajectory-based view

Recall from the introduction that coherent sets as defined in [Fro13, DJM16] rely on the addition of diffusion to the deterministic dynamics; this is so that large boundaries are penalised through greater diffusive mixing. The process we will consider, which gives rise

to the Fokker–Planck (Kolmogorov forward) equation (1), is a time-inhomogeneous Itô diffusion process governed by the SDE

$$d\mathbf{x}_t = v(t, \mathbf{x}_t)dt + \varepsilon d\mathbf{b}_t, \quad (8)$$

where v is a sufficiently smooth d -dimensional divergence-free velocity field, $\varepsilon \geq 0$, and \mathbf{b}_t is a d -dimensional standard Wiener process. We assume \mathbf{x}_0 to be uniformly distributed, thus the process is stationary. Let the set of initial conditions, M , be a smooth flat d -dimensional manifold (usually a subset of \mathbb{R}^d with smooth boundary, a cylinder, or a torus) equipped with a (Euclidean) metric. If M has a boundary, (8) is equipped with reflecting boundary conditions on $\bigcup_{t \in [0, \tau]} \{t\} \times \partial(\phi_t(M))$ in space-time; maintaining stationarity. Note that the domain for the SDE co-evolves with the deterministic flow driven by the velocity field v , thus (8) lives on \mathbb{M}_1 .

Next we will view the temporal component t of (8) as an independent variable θ undergoing diffusion. Since this diffusion can move in both directions along a line, the time component of (8) will also evolve in positive and in negative directions. We assume that the temporal parameter θ performs a Brownian diffusion with reflecting boundary conditions on $[0, \tau]$ and constant diffusion coefficient $a > 0$:

$$d\theta_t = a d\mathbf{w}_t, \quad (9)$$

where \mathbf{w}_t is a standard one-dimensional Wiener process independent of \mathbf{b}_t . Equation (8) now becomes

$$d\mathbf{x}_t = a v(\theta_t, \mathbf{x}_t) \circ d\mathbf{w}_t + \varepsilon d\mathbf{b}_t, \quad (10)$$

where we now assumed the Stratonovich form, and will explain the reason for this below. We call this the *time-diffused process*. An equivalent form of the system of equations (9) and (10) can be written in \mathbb{R}^{d+1} as

$$d\mathbf{X}_t = \underbrace{\begin{pmatrix} a & 0^\top \\ a v(\theta_t, \mathbf{x}_t) & \varepsilon \text{Id}_{d \times d} \end{pmatrix}}_{=: \sigma(\mathbf{X}_t)} \circ d\mathbf{B}_t, \quad (11)$$

with

$$\mathbf{X}_t = \begin{pmatrix} \theta_t \\ \mathbf{x}_t \end{pmatrix}, \quad \mathbf{B}_t = \begin{pmatrix} \mathbf{w}_t \\ \mathbf{b}_t \end{pmatrix}.$$

Note that we use upper case letters to denote the time-augmented version of a variable. The main reason for the the Stratonovich interpretation in (10) and (11) is satisfaction of the chain rule, which is important in the first of the following two situations.

1. **Spatial deterministic limit:** $\varepsilon \rightarrow 0$. When $\varepsilon = 0$, there is no spatial noise, and so we would like the paths of (9)–(10) to *stay on trajectories* $\{(\theta, x_\theta) \mid \theta \in [0, \tau]\}$ of the deterministic ODE $\dot{x}_t = v(t, x_t)$. This is only guaranteed in the Stratonovich case [Hsu02, Proposition 1.2.8]. For a deterministic trajectory $(x_r)_{r \in [0, \tau]}$ parametrized by the random time parameter $r = \theta_t$, i.e., $\mathbf{z}_t := x_{\theta_t}$, we have

$$d\mathbf{z}_t = \frac{d}{dr} x_r \Big|_{r=\theta_t} \circ d\theta_t = v(\theta_t, \mathbf{z}_t) \circ (a d\mathbf{w}_t),$$

exactly (10) for $\varepsilon = 0$. Since the solution of (10) is unique (for a fixed realization of the process \mathbf{w}_t), $\mathbf{z}_t = x_{\theta_t}$ is this solution, and it clearly evolves along trajectories of the deterministic ODE.

2. **High temporal diffusion limit:** $a \rightarrow \infty$. A similar situation occurs if instead we fix $\varepsilon > 0$ and let the temporal diffusion coefficient a increase to infinity. In the Lagrangian frame \mathbb{M}_0 , the drift v in (10) is zero and Φ -pullbacks of the paths of \mathbf{X}_t become increasingly aligned with the time axis because the stochastic variation in the temporal component (controlled by a) dominates the stochastic variation in space (controlled by ε). In \mathbb{M}_1 , paths of \mathbf{X}_t therefore align with trajectories $\{(\theta, x_\theta) \mid \theta \in [0, \tau]\}$ with overwhelming probability in the $a \rightarrow \infty$ limit.

We will return to these situations in section 3.2, showing that they are effectively equivalent. In the situation where ε is large relative to a , the stochastic trajectories of (9)–(10) may significantly deviate from the deterministic trajectories, and this deviation will be crucial for relaxing the strictly material nature of FTCSs.

2.2. The time-diffused process and Brownian motion on \mathbb{M}_1

So far the processes \mathbf{b}_t and \mathbf{w}_t have been standard Wiener processes with respect to the Euclidean metric on $\phi_t(M)$, $0 \leq t \leq \tau$. Next we wish to interpret certain processes as (standard) Brownian motions with respect to a suitable (Riemannian) metric. For clarity, in these cases the metric (or the entire Riemannian manifold) will be explicitly stated. For this interpretation, we will use the fact that standard Brownian motion on a Riemannian manifold is (up to equivalence in law) given by its generator, which is the *one half* Laplace–Beltrami operator on the manifold [Hsu02, Chapter 3].

The process we would like to interpret as Brownian motion on a Riemannian manifold is \mathbf{X}_t , governed by the SDE (11). This augmented process lives on the augmented spacetime manifold \mathbb{M}_1 , defined by (4), and so we need to find a metric on \mathbb{M}_1 whose Laplace–Beltrami operator is the generator of this process. One reason for seeking this connection is to link our construction to geometric characterisations of coherent sets [Fro15]. A second reason is to further develop our formalism for coherent sets when the “materialness requirement” that all previous work relied on is relaxed. In our derivation, we will rely on linking a SDE—in law—to a Riemannian metric via the associated Fokker–Planck (Kolmogorov forward) equation and the Laplace–Beltrami operator. Here, the Fokker–Planck equation identifies a SDE uniquely only in law, not pathwise.

We recall that we assumed v to be divergence-free. We expect to be able to derive similar results to those below for the non-divergence-free case as well, however, at the cost of more technical exposition that would obscure the main points. With this assumption, using the definition of σ in (11), it is immediate¹ that $\nabla \cdot \sigma^\top := \nabla_{(\theta, x)} \cdot \sigma^\top \equiv 0$. Thus, the Itô form of (11) is² [Pav14, p. 63]

$$d\mathbf{X}_t = \frac{1}{2} \nabla \cdot \Sigma_1(\mathbf{X}_t) dt + \sigma(\mathbf{X}_t) d\mathbf{B}_t, \quad \text{where } \Sigma_1 = \sigma \sigma^\top, \quad (12)$$

¹The divergence operator is applied to a matrix row-wise.

²The Itô form of a Stratonovich SDE $d\mathbf{z}_t = v dt + \sigma \circ d\mathbf{w}_t$, where the drift v and the diffusion coefficient σ depend on time and space as well, is given by $d\mathbf{z}_t = (v + \frac{1}{2}(\nabla \cdot (\sigma \sigma^\top) - \sigma \nabla \cdot \sigma^\top)) dt + \sigma d\mathbf{w}_t$.

and the drift is $\frac{1}{2}\nabla \cdot \Sigma_1$. More importantly, the corresponding Fokker–Planck operator (forward Kolmogorov generator, acting on the usual Sobolev space H^2 , or in a variational characterisation described by a bilinear form on H^1) $\mathcal{L}_1^* : H^2(\mathbb{M}_1) \rightarrow L^2(\mathbb{M}_1)$, becomes³ [Pav14, Proposition 3.4]

$$\begin{aligned}\mathcal{L}_1^* F &= \nabla \cdot \left(-\frac{1}{2}\nabla \cdot \Sigma_1 F + \frac{1}{2}\nabla \cdot (\Sigma_1 F) \right) = \frac{1}{2}\nabla \cdot \left(-\nabla \cdot \Sigma_1 F + \nabla \cdot \Sigma_1 F + \Sigma_1 \nabla F \right) \\ &= \frac{1}{2}\nabla \cdot (\Sigma_1 \nabla F),\end{aligned}\tag{13}$$

with homogeneous Neumann boundary conditions both in space and time, corresponding to the reflecting boundary conditions for the SDE. The subscripts in Σ_1 and \mathcal{L}_1 indicate that these objects are naturally connected to \mathbb{M}_1 , and we use \mathcal{L}_1^* to denote the Fokker–Planck operator, despite the above operator being self-adjoint.

Proposition 1. *We have $\mathcal{L}_1^* = \frac{1}{2}\Delta_{G_1}$, where Δ_{G_1} is the Laplace–Beltrami operator corresponding to the metric G_1 given by the metric tensor*

$$G_1(\theta, x) = \Sigma_1(\theta, x)^{-1} = \begin{pmatrix} \frac{1}{a^2} + \frac{1}{\varepsilon^2}v(\theta, x)^\top v(\theta, x) & -\frac{1}{\varepsilon^2}v(\theta, x)^\top \\ -\frac{1}{\varepsilon^2}v(\theta, x) & \frac{1}{\varepsilon^2}\text{Id} \end{pmatrix}\tag{14}$$

in the Euclidean coordinates (θ, x) on \mathbb{M}_1 .

Proof. For a general metric g the Laplace–Beltrami operator is given by

$$\Delta_g f = \frac{1}{\sqrt{|g|}} \nabla \cdot \left(\sqrt{|g|} g^{-1} \nabla f \right),\tag{15}$$

where $|g|$ denotes the absolute value of the determinant of the metric tensor, the latter also denoted by g . In our particular situation, from (11) we have that $\det \sigma = a\varepsilon^d$, and hence $\det \Sigma_1 = (\det \sigma)^2 = a^2\varepsilon^{2d}$. Since the determinant of Σ_1 is constant, the volume distortion $\sqrt{|G_1|}$ cancels out and we obtain $\Delta_{G_1} F = \nabla \cdot (\Sigma_1 \nabla F)$ for a sufficiently smooth function $F : \mathbb{M}_1 \rightarrow \mathbb{R}$. A comparison with (13) implies the claim. The explicit form can be seen from $\Sigma_1^{-1} = \sigma^{-\top} \sigma^{-1}$, (11), and the identity

$$\begin{pmatrix} a & 0^\top \\ v & J \end{pmatrix}^{-1} = \begin{pmatrix} a^{-1} & 0^\top \\ -a^{-1}J^{-1}v & J^{-1} \end{pmatrix} \quad \text{for } a \in \mathbb{R} \setminus \{0\}, v \in \mathbb{R}^d, J \in \mathbb{R}^{d \times d} \text{ invertible.}\tag{16}$$

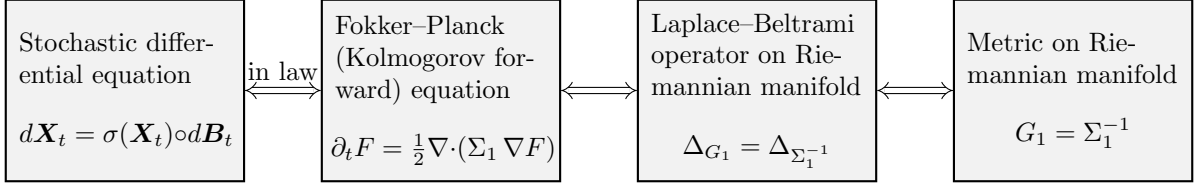
□

The following is an immediate consequence of Proposition 1. Note that we interchangeably use the metric and its metric tensor expressed in Euclidean coordinates.

Corollary 2. *The generator of the process \mathbf{X}_t governed by (11) on \mathbb{M}_1 is given by $\mathcal{L}_1^* = \frac{1}{2}\Delta_{\Sigma_1^{-1}}$. Hence, \mathbf{X}_t is equivalent in law to the standard Brownian motion on the Riemannian manifold (\mathbb{M}_1, G_1) , where $G_1 = \Sigma_1^{-1}$.*

³From [Pav14, pp. 69 & 71] we have that an Itô SDE $d\mathbf{z}_t = v dt + \sigma d\mathbf{w}_t$ has the forward generator $\mathcal{L}^* f = \nabla \cdot (-vf + \frac{1}{2}\nabla \cdot (\sigma\sigma^\top f)) = \nabla \cdot \left(\left(-v + \frac{1}{2}\nabla \cdot (\sigma\sigma^\top) \right) f + \frac{1}{2}\sigma\sigma^\top \nabla f \right)$.

The following diagram summarises the one-to-one relationships we have used between objects associated with \mathbb{M}_1 , where the first one is merely in law:



Next we will use the fact that these relationships remain valid under pullback of the respective objects from \mathbb{M}_1 to \mathbb{M}_0 .

2.3. The pullback of the time-diffused process on \mathbb{M}_0

2.3.1. The pullback process

A deterministic trajectory $\{\phi_\theta(x)\}_{0 \leq \theta \leq \tau}$ initialized at x at time $\theta = 0$ lies on the curve $\{(\theta, \phi_\theta(x)) : 0 \leq \theta \leq \tau\} \subset \mathbb{M}_1$. As these curves are parametrized by their initial conditions in M , we will now pull back the time-diffused process (11) to the manifold \mathbb{M}_0 consisting of time and initial-condition pairs. This amounts to describing the process in what is often called the Lagrangian frame.

Recall from (5) the canonical mapping $\Phi : (\theta, x) \mapsto (\theta, \phi_\theta(x))$ from \mathbb{M}_0 to the trajectory manifold \mathbb{M}_1 . It has the Jacobian matrix

$$D\Phi(\theta, x) = \begin{pmatrix} 1 & 0^\top \\ v(\theta, \phi_\theta x) & J_\theta(x) \end{pmatrix}, \quad (17)$$

where $J_\theta(x) := \partial_x(\phi_\theta(x))$ is a shorthand for the Jacobian matrix of the time- θ -flow map of the deterministic system (i.e., the ODE (8) with $\varepsilon = 0$).

If $\mathbf{Y}_t := \Phi^{-1}(\mathbf{X}_t)$ denotes the pulled-back time-diffused process (11), then this satisfies the SDE $d\mathbf{Y}_t = (D\Phi^{-1})(\mathbf{X}_t)\sigma(\mathbf{X}_t) \circ d\mathbf{B}_t$, since SDEs in Stratonovich form obey the chain rule [Pav14, Prop. 3.4]. Using that $(D\Phi^{-1})(\Phi(Y)) = (D\Phi(Y))^{-1}$ and the identity (16) we obtain that

$$\begin{aligned} d\mathbf{Y}_t &= \begin{pmatrix} 1 & 0^\top \\ -J_{\theta_t}(\mathbf{y}_t)^{-1}v(\theta_t, \mathbf{x}_t) & J_{\theta_t}(\mathbf{y}_t)^{-1} \end{pmatrix} \begin{pmatrix} a & 0^\top \\ av(\theta_t, \mathbf{x}_t) & \varepsilon \text{Id} \end{pmatrix} \circ d\mathbf{B}_t \\ &= \begin{pmatrix} a & 0^\top \\ 0 & \varepsilon J_{\theta_t}(\mathbf{y}_t)^{-1} \end{pmatrix} \circ d\mathbf{B}_t, \end{aligned} \quad (18)$$

with

$$\mathbf{Y}_t = \begin{pmatrix} \theta_t \\ \mathbf{y}_t \end{pmatrix}, \quad \mathbf{B}_t = \begin{pmatrix} \mathbf{w}_t \\ \mathbf{b}_t \end{pmatrix}.$$

Hence, the pullback to \mathbb{M}_0 (block-)diagonalises the diffusion coefficient (matrix) of the time-diffused process, so that the independent noise processes \mathbf{w}_t and \mathbf{b}_t only influence the time and space parameters θ and y , respectively. This is reflected in the Laplace–Beltrami operator (20) below, where there are no mixed time and space-derivatives.

The dynamic interpretation of this process is straightforward. If we pull back the time-diffused process to \mathbb{M}_0 , where deterministic trajectories are straight lines, then

- the noise \mathbf{w}_t drives the diffusion *along* a single trajectory,
- \mathbf{b}_t drives the diffusion *between* trajectories, and
- the influence of the nonlinear dynamics on the diffusion between trajectories is encoded in the appearance of the Jacobian matrices J_θ .

2.3.2. The pullback metric

To obtain a characterisation of the pulled-back process \mathbf{Y}_t in law, we will now pull back the metric G_1 to \mathbb{M}_0 and consider the Brownian motion it generates thereon.

We briefly recall some general facts about pullback metrics. Let $\Psi : M \rightarrow N$ be a diffeomorphism between two smooth manifolds, and endow N with a metric n . The pullback metric Ψ^*n on M is defined by $(\Psi^*n)_x(v, w) = n_{\Psi(x)}(D\Psi v, D\Psi w)$, the subscripts referring to the point at which the metric is evaluated. If in local coordinates at a point $\Psi(x) \in N$ the metric n is expressed by the metric tensor \mathcal{N} , then the metric Ψ^*n is expressed in local coordinates at $x \in M$ by $(D\Psi(x))^\top \mathcal{N} D\Psi(x)$.

Note that for both \mathbb{M}_0 and \mathbb{M}_1 , local coordinates coincide with the global Euclidean coordinates, and recall that $G_1 = \Sigma_1^{-1}$. With $D\Phi$ in (17), we obtain by Proposition 1 and the local coordinate expression for the pullback above that the pullback metric $G_0 := \Phi^*G_1$ with inverse metric tensor Σ_0 satisfies

$$\begin{aligned} \Sigma_0(\theta, y)^{-1} &:= G_0(\theta, y) = (\Phi^*G_1)(\theta, y) = D\Phi(\theta, y)^\top \Sigma_1(\theta, \phi_\theta y)^{-1} D\Phi(\theta, y) \\ &= \begin{pmatrix} \frac{1}{a^2} & 0^\top \\ 0 & \frac{1}{\varepsilon^2} J_\theta(y)^\top J_\theta(y) \end{pmatrix}. \end{aligned} \quad (19)$$

We summarize this discussion in Figure 2.

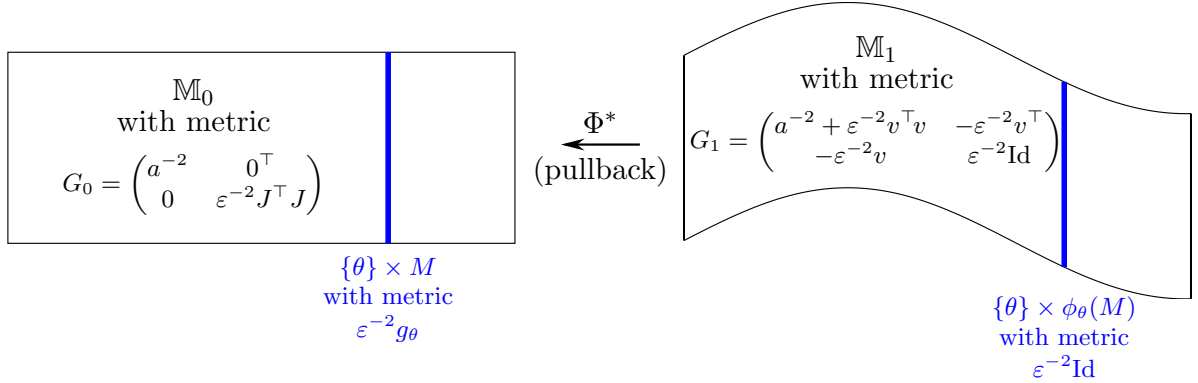


Figure 2.: The metric G_1 associated with the time-diffused process on \mathbb{M}_1 and its pullback G_0 on \mathbb{M}_0 . The induced metrics on the θ -time-slices of the manifolds (blue vertical lines) are $\frac{1}{\varepsilon^2} g_\theta$ and $\frac{1}{\varepsilon^2} \text{Id}$, respectively. Note that the pullback decouples the metric—and hence the diffusion coefficient of the process—into purely temporal and spatial components.

Recall from the introduction that g_θ is the pullback of the Euclidean metric by ϕ_θ to M , namely ϕ_θ^*e . In Euclidean coordinates we have $g_\theta(y) = J_\theta(y)^\top J_\theta(y)$. Thus, denoting by

Δ_{g_θ} the Laplace–Beltrami operator with respect to the metric g_θ and using (15) and (19), the Laplace–Beltrami operator of the Riemannian manifold (\mathbb{M}_0, G_0) is given by

$$\Delta_{G_0} F(\theta, \cdot) = a^2 \partial_\theta^2 F(\theta, \cdot) + \varepsilon^2 \Delta_{g_\theta} F(\theta, \cdot) =: 2\mathcal{L}_0^* F(\theta, \cdot), \quad (20)$$

because $\det J_\theta \equiv 1$. Again, the associated boundary conditions are homogeneous Neumann. Note that in analogy to the previous subsection and Proposition 1 we denote in (20) the generator of the standard Brownian motion on (\mathbb{M}_0, G_0) by $\mathcal{L}_0^* = \frac{1}{2} \Delta_{G_0}$. Similarly to Corollary 2 we obtain:

Corollary 3. *The generator of the pulled-back time-diffused process \mathbf{Y}_t governed by (18) on \mathbb{M}_0 is given by $\mathcal{L}_0^* = \frac{1}{2} \Delta_{\Sigma_0^{-1}}$. Hence, \mathbf{Y}_t is equivalent in law to the standard Brownian motion on the Riemannian manifold (\mathbb{M}_0, G_0) , where $G_0 = \Sigma_0^{-1}$.*

Remark 4. *The pullback to the initial time $\theta = 0$ is simply for convenience. Let $\phi_{\vartheta, \theta}$ denote the flow of the ODE $\dot{x}_t = v(t, x_t)$ from time ϑ to time θ , i.e., $\phi_{\vartheta, \theta} = \phi_\theta \circ \phi_\vartheta^{-1}$. We note that we could use the manifold of states at any time $\vartheta \in [0, \tau]$, $M_\vartheta := \phi_{0, \vartheta}(M)$, to define $\mathbb{M}_\vartheta := \bigcup_{\theta \in [0, \tau]} \{\theta\} \times M_\vartheta$, and pull back the metric G_1 to \mathbb{M}_ϑ using*

$$\Phi_\vartheta : (\theta, z) \mapsto (\theta, \phi_{\vartheta, \theta}(z)).$$

Then the formulas (19) and (20) hold with $J_\theta = \partial_z(\phi_{\vartheta, \theta}(z))$ and g_θ modified accordingly. In particular, the spectra of the Laplace–Beltrami operators Δ_{G_0} and $\Delta_{\Phi_\vartheta^ G_1}$ coincide, and the associated eigenfunctions can be obtained by coordinate transformation with $z = \phi_{0, \vartheta}(y)$.*

Next we will see how the time-diffused process connects to the currently established notion of coherent sets, and how it extends this notion to semi-material coherent sets.

3. The inflated dynamic Laplace operator and connections to the dynamic Laplace operator

In the following, the dependence of the metric G_0 and of related objects on the parameter a is going to be of central interest. Hence, we will explicitly denote this dependence by writing $G_{0,a}$. The parameter ε is the strength of the spatial diffusion; see (1) and (8). As discussed in the introduction, this is usually taken to be small [Fro13], or in the vanishing limit when constructing the dynamic Laplace operator Δ^D [Fro15]. Our inflated dynamic Laplace operator (20) corresponding to the SDE (18) has the additional parameter a . In this section we will show that for any fixed $\varepsilon > 0$, in the limit as $a \rightarrow \infty$ the properties of $\Delta_{G_{0,a}}$ mimic those of Δ^D . Therefore, for simplicity from now on we almost exclusively set $\varepsilon = 1$ and retain only the parameter a .

3.1. The dynamic Laplace operator

The dynamic Laplace operator Δ^D in (2) arose from the desire to find (material) coherent sets [Fro15] using purely geometric constructions. Coherent sets $A \subset M$ maximally

inhibit mixing due to their boundary size remaining small under evolution by ϕ_t relative to enclosed volume. An evolving boundary remaining small is a common measure of advective mixing [Pie91, Ott89] because in the presence of small diffusion the dispersion of mass outside $\phi_t(A)$ is proportional to the boundary size of $\phi_t(A)$. This small evolving boundary geometry has been encoded in statements for a new dynamic isoperimetric theory [Fro15]. In particular, the classical Cheeger inequality, which relates the geometry of a Riemannian manifold to the first nontrivial eigenvalue of its Laplace–Beltrami operator [Che70], and the classical Federer–Fleming theorem [FF60], which equates the Cheeger constant and the Sobolev constant, were each extended to the dynamic situation in [Fro15, FK20] and applied to defining and discovering finite-time coherent sets. One aspect we will repeatedly use is that leading nontrivial eigenvalues of both Δ^D and $\Delta_{G_{0,a}}$ are strongly connected with coherence in our dynamical system. Links between these dynamic isoperimetric quantities and the spectrum of Δ^D , and their counterparts in our time-expanded geometry will be taken up in this section.

3.2. The dynamic Laplace operator as the averaging-limit

We show that the dynamic Laplace operator Δ^D arises from $\Delta_{G_{0,a}}$ in the limit of infinite temporal diffusion. The intuition from the metric point of view is to note that as $a \rightarrow \infty$ the distance according to $G_{0,a}$ between two points (θ_1, x) and (θ_2, x) on the same deterministic trajectory in \mathbb{M}_0 goes to zero; see the left-hand image in Figure 2. Therefore, in the $a \rightarrow \infty$ limit the importance of temporal displacements vanishes and all relevant information is captured by the temporal average (2) defining Δ^D .

To demonstrate this convergence formally we consider the statistical properties of the process (18) in the $a \rightarrow \infty$ limit. To this end recall that the generator \mathcal{L}_0^* of (18) is the half Laplace–Beltrami operator $\Delta_{G_{0,a}}$ from (20). Using Footnote 3, we see that an SDE in Itô form having generator \mathcal{L}_0^* (20) is given by

$$\begin{aligned} d\boldsymbol{\theta}_t &= a d\mathbf{w}_t, \\ d\mathbf{y}_t &= \frac{1}{2} \nabla_{\mathbf{y}} \cdot (J_{\boldsymbol{\theta}_t}^{-1} J_{\boldsymbol{\theta}_t}^{-\top}) (\mathbf{y}_t) dt + J_{\boldsymbol{\theta}_t}(\mathbf{y}_t)^{-1} d\mathbf{b}_t, \end{aligned} \quad (21)$$

The temporal process $\boldsymbol{\theta}_t$ is much faster than the spatial process when $a \gg 1$, and thus the results of averaging can be applied. More precisely, it follows from [PS08, Remark 10.2, Section 10.7, Section 17.4] that the slow process \mathbf{y}_t converges weakly (i.e., in distribution) to the averaged process

$$d\bar{\mathbf{y}}_t = \bar{v}(\bar{\mathbf{y}}_t) dt + \bar{\sigma}(\bar{\mathbf{y}}_t) d\mathbf{b}_t, \quad (22)$$

as $a \rightarrow \infty$, where $\bar{v} := \frac{1}{2} \nabla \cdot \bar{\Sigma}$, $\bar{\sigma}$ is any matrix-valued function satisfying $\bar{\sigma} \bar{\sigma}^\top = \bar{\Sigma}$, and

$$\bar{\Sigma}(y) := \frac{1}{\tau} \int_0^\tau (J_\theta^{-1} J_\theta^{-\top}) (y) d\theta. \quad (23)$$

Note that $\bar{\Sigma}$ is symmetric and positive definite as an integral of symmetric positive-definite matrices, thus one can find a (nonunique) $\bar{\sigma}$ such that $\bar{\sigma} \bar{\sigma}^\top = \bar{\Sigma}$.

We see by Footnote 3 that the limiting slow process $\bar{\mathbf{y}}_t$ has the forward generator $\bar{\mathcal{L}}^* : H^2(M) \rightarrow L^2(M)$ given by

$$\bar{\mathcal{L}}^* f = \nabla \cdot \left((-\bar{v} + \underbrace{\frac{1}{2} \nabla \cdot \bar{\Sigma}}_{=\bar{v}}) f + \frac{1}{2} \bar{\Sigma} \nabla f \right) = \frac{1}{2} \nabla \cdot (\bar{\Sigma} \nabla f),$$

which does not depend on the particular choice of $\bar{\sigma}$. Recall that $g_\theta(y)$ has matrix representation $J_\theta^{-1} J_\theta^{-\top}(y)$, so by (15) we have $\Delta_{g_\theta} f = \nabla \cdot (J_\theta^{-1} J_\theta^{-\top} \nabla f)$. We can now immediately see from (2) and (23) that $\frac{1}{2} \Delta^D = \bar{\mathcal{L}}^*$. In other words, the dynamic Laplace operator is (twice) the generator of the dominant spatial process \mathbf{y}_t in (21) as $a \rightarrow \infty$. The expression for $\bar{\Sigma}$ appears as a harmonic mean of the metrics g_t in [KK20]; here it arises naturally in the limit of speeding up time in the temporal diffusion.

Let us briefly re-introduce the parameter $\varepsilon > 0$ into (21) to compare the two situations as at the end of section 2.1, where we first considered the limit $\varepsilon \rightarrow 0$ for fixed $a > 0$, and then the limit $a \rightarrow \infty$ for fixed $\varepsilon > 0$:

$$\begin{aligned} d\boldsymbol{\theta}_t &= a d\mathbf{w}_t, \\ d\mathbf{y}_t &= \frac{\varepsilon^2}{2} \nabla_y \cdot (J_{\boldsymbol{\theta}_t}^{-1} J_{\boldsymbol{\theta}_t}^{-\top}) (\mathbf{y}_t) dt + \varepsilon J_{\boldsymbol{\theta}_t} (\mathbf{y}_t)^{-1} d\mathbf{b}_t. \end{aligned}$$

This SDE is in law equivalent to the pullback of (11). The theory of averaging [PS08] allows one to draw an equivalence between these two limits: the limiting evolution of the (slow) y -component is in law governed by (22) in both cases, for $\varepsilon \rightarrow 0$ on the timescale $t = \mathcal{O}(\varepsilon^{-2})$, and for $a \rightarrow \infty$ on the timescale $t = \mathcal{O}(1)$, respectively. Hence, the two situations can be transformed into one another via a suitable rescaling of time. We remark that Δ^D can also be obtained [KS21] as the leading-order term for $\varepsilon \rightarrow 0$ in the Fokker–Planck equation (1) viewed in Lagrangian coordinates [Thi03] (i.e. where all times $t \in [0, \tau]$ are pulled back to $t = 0$), using operator-averaging techniques adapted from [Kro91].

3.3. Cheeger and Sobolev inequalities

We now begin to analyse the Laplace–Beltrami operator $\Delta_{G_{0,a}}$ and the associated metric $G_{0,a}$. In the sequel we rarely consider time-evolution of the associated augmented process (18), and so we revert to denoting the temporal coordinate of the augmented manifold \mathbb{M}_0 by t (instead of θ), to stress its temporal character. The purpose of this section is to link the Cheeger and Sobolev constants of our time-augmented manifold with the dynamic Cheeger and dynamic Sobolev constants of [Fro15, FK20].

Suppose that $\Gamma \subset M$ is a co-dimension 1 C^∞ surface disconnecting M into the disjoint union $M = A_1 \cup \Gamma \cup A_2$, with A_1, A_2 connected submanifolds. Let $\iota : \Gamma \hookrightarrow M$ denote the inclusion map, $\iota^* g_t$ the induced metric on Γ arising from g_t , and $V_{\iota^* g_t}$ the corresponding volume form on Γ . Recall the dynamic Cheeger constant [Fro15, equation (20)] or [FK20, equations (3.6)–(3.7)], which we can write as

$$h^D := \inf_{\Gamma} \frac{\frac{1}{\tau} \int_0^\tau V_{\iota^* g_t}(\Gamma) dt}{\min\{\ell(A_1), \ell(A_2)\}}. \quad (24)$$

In the expression (24), we select a Γ disconnecting M and follow its forward evolution under the nonlinear dynamics ϕ_t . We wish to find the initial disconnecter Γ whose average evolved size is least, relative to the volumes of the two connected components of M , as this represents the potential boundary of a finite-time coherent set. The dynamic Cheeger inequality [Fro15, FK20] states that

$$h^D \leq 2\sqrt{-\lambda_2^D}, \quad (25)$$

where λ_2^D is the first nontrivial eigenvalue of Δ^D .

To motivate the next construction, we note the evolution of any surface Γ in (24) that disconnects M will trace out a surface $\mathbb{T}' := \bigcup_{t \in [0, \tau]} (\{t\} \times \phi_t \Gamma)$ that disconnects our “forward-evolved” time-augmented manifold \mathbb{M}_1 . With Φ from (5) we may pull back such a traced-out surface to obtain a surface $\mathbb{T} = \Phi^{-1} \mathbb{T}'$ disconnecting \mathbb{M}_0 . Of course by the construction of \mathbb{T}' and the definition of Φ , the surface \mathbb{T} has a constant section on each time fibre, namely Γ . That is, $\mathbb{T} \cap (\{t\} \times M) = \Gamma$ for each $t \in [0, \tau]$. In summary, the minimising disconnecter Γ from (24) provides a particular “constant-in-time” disconnecter \mathbb{T} of our augmented Riemannian manifold (\mathbb{M}_0, G_a) . However, we may also consider more general disconnectors of (\mathbb{M}_0, G_a) , which would represent a relaxation of materialness in the standard definition of coherent sets [Fro15]. This can be accomplished using the standard Cheeger constant for the manifold (\mathbb{M}_0, G_a) for suitable a . Let $\mathfrak{i} : \mathbb{T} \hookrightarrow \mathbb{M}_0$ denote the inclusion map, $\mathfrak{i}^* G_{0,a}$ the induced metric on \mathbb{T} , and $V_{\mathfrak{i}^* G_{0,a}}$ the corresponding volume form on \mathbb{T} . Set

$$H_a := \inf_{\mathbb{T}} \frac{V_{\mathfrak{i}^* G_{0,a}}(\mathbb{T})}{\min\{V_{G_{0,a}}(\mathbb{A}_1), V_{G_{0,a}}(\mathbb{A}_2)\}}, \quad (26)$$

where $\mathbb{T} \subset \mathbb{M}_0$ is a co-dimension 1 C^∞ surface disconnecting \mathbb{M}_0 into the disjoint union $\mathbb{M}_0 = \mathbb{A}_1 \cup \mathbb{T} \cup \mathbb{A}_2$, with $\mathbb{A}_1, \mathbb{A}_2$ connected submanifolds. Such a \mathbb{T} is a potential boundary of a semi-material coherent set with parameter a . The standard Cheeger inequality in this situation reads

$$H_a \leq 2\sqrt{-\Lambda_2}, \quad (27)$$

where Λ_2 is the first nontrivial eigenvalue of $\Delta_{G_{0,a}}$. We will shortly address the relationship between h^D and H_a , and the behaviour of H_a with increasing a .

Similarly, we recall the dynamic Sobolev constant [Fro15, equation (21)] or [FK20, §3.2]:

$$s^D := \inf_{f \in C^\infty(M)} \frac{\frac{1}{\tau} \int_0^\tau \int_M \|\nabla_{g_t} f\|_{g_t} d\ell dt}{\inf_{\alpha \in \mathbb{R}} \int_M |f - \alpha| d\ell}, \quad (28)$$

where $\nabla_g f$ is the unique vector field on M satisfying

$$g(\nabla_g f, w) = w(f) \quad (29)$$

for all vector fields $w : M \rightarrow \mathbb{R}^d$, and $w(f)$ is the Lie derivative of $f : M \rightarrow \mathbb{R}$. Let us unpick these definitions in coordinates. In coordinates, $w(f)(x) = \sum_{i=1}^d w_i(x) \cdot \frac{\partial f}{\partial x_i}(x)$, the directional derivative at x of f in the direction $w(x)$; we will denote by $[\partial f(x)]$ the

matrix of partial derivatives evaluated at x . The expression $g(\nabla_g f, w)(x)$ in coordinates is $\nabla_g f(x)^\top [g(x)] w(x)$, where $[g(x)]$ is the coordinate matrix⁴ representation of g at $x \in M$. Thus, in coordinates (29) becomes $\nabla_g f(x)^\top [g(x)] w(x) = [\partial f(x)]^\top w(x)$. Since this holds for all vector fields w we have

$$\nabla_g f(x) = [g(x)]^{-1} \partial f(x), \quad (30)$$

using symmetry and invertibility of $[g(x)]$. Finally we have

$$\begin{aligned} \|\nabla_g f(x)\|_g^2 &= g(\nabla_g f(x), \nabla_g f(x)) \\ &= \nabla_g f(x)^\top [g(x)] \nabla_g f(x) \\ &= [\partial f(x)]^\top [g(x)^{-1}] [\partial f(x)] \quad \text{by (30)}. \end{aligned} \quad (31)$$

We may also define a Sobolev constant for the Riemannian manifold $(\mathbb{M}_0, G_{0,a})$ in the usual way:

$$S_a := \inf_{F \in C^\infty(\mathbb{M}_0)} \frac{\int_{\mathbb{M}_0} \|\nabla_{G_{0,a}} F\|_{G_{0,a}} dV_{G_{0,a}}}{\inf_{\alpha \in \mathbb{R}} \int_{\mathbb{M}_0} |F - \alpha| dV_{G_{0,a}}}, \quad (32)$$

where $V_{G_{0,a}}$ denotes the volume measure with respect to $G_{0,a}$. Note that by (19) we have $dV_{G_{0,a}} = \frac{1}{a} dt d\ell$.

Proposition 5. *One has*

1. $H_a = S_a \leq s^D = h^D$ for all $a \geq 0$,
2. H_a and S_a are nondecreasing in $a \geq 0$.

Proof.

1. The fact that $H_a = S_a$ follows from the Federer–Fleming Theorem (e.g. [Cha84, p. 131]). The fact that $s^D = h^D$ follows from the dynamic Federer–Fleming Theorem (Theorem 3.1 [Fro15], Theorem 3.3 [FK20]).

We now treat the inequality $S_a \leq s^D$. For $f \in C^\infty(M)$, denote by $s^D(f)$ the infimum of (28). For $\varepsilon > 0$ let $f_\varepsilon \in C^\infty(M)$ be such that $s^D(f_\varepsilon) \leq s^D + \varepsilon$. Define $F_\varepsilon \in C^\infty(\mathbb{M}_0)$ by $F_\varepsilon(t, x) = f_\varepsilon(x)$ for $(t, x) \in \mathbb{M}_0$. Then

$$\begin{aligned} S_a(F_\varepsilon) &:= \frac{\int_{\mathbb{M}_0} \|\nabla_{G_{0,a}} F_\varepsilon\|_{G_{0,a}} dV_{G_{0,a}}}{\inf_{\alpha \in \mathbb{R}} \int_{\mathbb{M}_0} |F_\varepsilon - \alpha| dV_{G_{0,a}}} \\ &= \frac{\int_{[0,\tau]} \int_M (a^2 (\partial_t F_\varepsilon)^2 + [\partial_x F_\varepsilon]^\top (J_t^\top J_t)^{-1} [\partial_x F_\varepsilon])^{1/2} \frac{1}{a} dt d\ell}{\inf_{\alpha \in \mathbb{R}} \int_{[0,\tau]} \int_M |F_\varepsilon - \alpha| \frac{1}{a} dt d\ell} \\ &= \frac{\frac{1}{\tau} \int_{[0,\tau]} \int_M ([\partial_x F_\varepsilon]^\top (J_t^\top J_t)^{-1} [\partial_x F_\varepsilon])^{1/2} dt d\ell}{\frac{1}{\tau} \left(\int_{[0,\tau]} dt \right) \inf_{\alpha \in \mathbb{R}} \int_M |F_\varepsilon - \alpha| d\ell} \\ &= s^D(f_\varepsilon) \leq s^D + \varepsilon. \end{aligned}$$

Since ε was arbitrary, this implies $S_a = \inf_{F \in C^\infty(\mathbb{M}_0)} S_a(F) \leq s^D$.

⁴We will occasionally, but not universally, use square brackets around objects like ∂f , g , and G to emphasise that an object is to be interpreted as a matrix.

2. By (19), (31), and (32) one has

$$S_a = \inf_{F \in C^\infty(\mathbb{M}_0)} \frac{\int_{[0,\tau]} \int_M \left(a^2 (\partial_t F)^2 + [\partial_x F]^\top (J_t^\top J_t)^{-1} [\partial_x F] \right)^{1/2} \frac{1}{a} dt d\ell}{\inf_{\alpha \in \mathbb{R}} \int_{[0,\tau]} \int_M |F - \alpha| \frac{1}{a} dt d\ell}.$$

The result follows by noting that the integrand is nondecreasing in a . By the equality $H_a = S_a$ in Part 1 we obtain that H_a is nondecreasing in a . □

Remark 6. *The values H_a and S_a quantify the maximum level of coherence present: low $H_a = S_a$ indicates strong coherence. Proposition 5 says that increasing a leads to greater boundary lengths relative to volume on \mathbb{M}_0 and therefore lower coherence. Referring to (19), with increasing a , the numerator in (26) can be reduced by aligning the tangent spaces of Γ with the time axis (recall we are always working in \mathbb{M}_0). Thus, as one increases a , we expect the minimising Γ to become increasingly material; for example, in the lower left panel of Figure 1) the boundary of the pale blue set will become more horizontal. In summary, there is a trade-off between materiality and coherence, with the former increasing and the latter decreasing with increasing a .*

3.4. Spectrum and eigenfunctions

The spectrum of the dynamic Laplace operator and our proposed inflated dynamic Laplace operator characterises the strength of coherence and suggests natural numbers of coherent sets. Eigenvalues near to zero indicate the presence of strong coherence, and their corresponding eigenfunctions encode the location of coherent sets in the phase space.

3.4.1. Spectrum of the dynamic Laplace operator

We begin by recalling the variational characterisation of eigenvalues of the dynamic Laplacian and then link these to our inflated dynamic Laplace operator $\Delta_{G_{0,a}}$ on augmented space. We consider the dynamic Laplacian Δ^D as defined in [Fro15, equation (28)] and [FK20, equation (4.12)]. The boundary condition on M is the natural one for the dynamic Laplacian and corresponds to a “dynamic Neumann boundary condition”; see [Fro15, equation (30)] for an explicit representation.

By [Fro15, Theorem 4.1] and [Fro15, Remark 4.2] or [FK20, Theorem 4.4] and the discussion in [FK20, Section 4.2], the dynamic Laplacian has a countable discrete spectrum $0 = \lambda_1^D > \lambda_2^D \geq \lambda_3^D \cdots$ with the corresponding eigenfunctions denoted $\mathbf{1}_M \equiv f_1, f_2, \dots \in C^\infty(M)$. Let $S_0 = L^2(M)$ and for $k \geq 1$ let $S_k = \{f \in L^2(M) : \langle f, f_i \rangle = 0, 1 \leq i \leq k\}$. By [Fro15, equation (34)], one has the following variational representation of λ_k^D for $k \geq 1$:

$$\lambda_k^D = - \inf_{f: M \rightarrow \mathbb{R}, f \in S_{k-1}} \frac{\int_0^\tau \int_M \|\nabla_{g_t} f(x)\|_{g_t}^2 d\ell(x) dt}{\int_0^\tau \int_M f(x)^2 d\ell(x) dt}. \quad (33)$$

3.4.2. Spectrum of the inflated dynamic Laplace operator

We recall that $\Delta_{G_{0,a}}$ is equipped with homogeneous Neumann boundary conditions. By standard theory (e.g. [Lab15]) $\Delta_{G_{0,a}}$ has a discrete spectrum $0 = \Lambda_{1,a} > \Lambda_{2,a} \geq \Lambda_{3,a} \dots$ with eigenfunctions $\mathbf{1}_{\mathbb{M}_0} \equiv F_1, F_2, \dots \in C^\infty(\mathbb{M}_0)$. Some of the eigenfunctions are easily identifiable: for $k \geq 1$, the functions $F_k^{\text{temp}}(t, \cdot) := \cos(k\pi t/\tau)$ are clearly eigenfunctions with eigenvalue $\Lambda_{k,a}^{\text{temp}} := -(a\pi k/\tau)^2$. We call these eigenfunctions *temporal modes* or *temporal eigenfunctions* because they are constant in space and vary only in time. Define $W_0 = \{f \in L^2([0, \tau]) : \int_0^\tau f(t) dt = 0\}$ and $\mathbb{S}_0^{\text{temp}} := \{F = f\mathbf{1}_M : f \in W_0\} \subset L^2(\mathbb{M}_0)$, a subspace containing all temporal eigenfunctions.

The operator $\Delta_{G_{0,a}}$ is symmetric on its domain in $L^2(\mathbb{M}_0)$ and so its eigenfunctions are L^2 -orthogonal. Therefore if F is a non-temporal eigenfunction, then $F \perp \cos(k\pi t/\tau)$ for all $k \geq 1$. As the temporal eigenfunctions are dense in $\mathbb{S}_0^{\text{temp}}$, we have that F is orthogonal to every function in $\mathbb{S}_0^{\text{temp}}$, which implies⁵ that F has constant spatial means:

$$\int_M F(\cdot, x) d\ell(x) = \text{const a.e. on } [0, \tau]. \quad (34)$$

We denote the subspace of all such functions by $\mathbb{S}_0^{\text{spat}} = (\mathbb{S}_0^{\text{temp}})^\perp$. The non-temporal eigenfunctions will be called *spatial eigenfunctions* or *spatial modes*. In general, they will vary both in space and in time. The associated eigenvalues will be denoted by $0 = \Lambda_{1,a}^{\text{spat}} > \Lambda_{2,a}^{\text{spat}} \geq \Lambda_{3,a}^{\text{spat}} \dots$. The spatial and temporal eigenvalues partition the spectrum $\sigma(\Delta_{G_{0,a}})$. By (34), an eigenfunction F is a spatial mode if and only if its spatial mean $t \mapsto \int_M F(t, \cdot) d\ell$ is an a.e. constant function of time. We will later use this distinct behavior to numerically distinguish between temporal and spatial modes.

3.4.3. Behaviour of the spectrum of $\Delta_{G_{0,a}}$ with increasing a

We next address the behavior of the eigenvalues of $\Delta_{G_{0,a}}$ with increasing $a > 0$, linking them to the eigenvalues of the dynamic Laplace operator. Let $\mathbb{S}_0 = L^2(\mathbb{M}_0)$ and for $k \geq 1$ let $\mathbb{S}_k = \{F \in L^2(\mathbb{M}_0) : \langle F, F_i \rangle = 0, 1 \leq i \leq k\}$. For $k \geq 1$, one has the standard variational characterisation of eigenvalues of Laplace–Beltrami operators (recall that the volume form $V_{G_{0,a}}$ is given by $dV_{G_{0,a}} = \frac{1}{a}d\ell$)

$$\Lambda_{k,a} = - \inf_{F: \mathbb{M}_0 \rightarrow \mathbb{R}, F \in \mathbb{S}_{k-1}} \frac{\int_{\mathbb{M}_0} \|\nabla_{G_{0,a}} F\|_{G_{0,a}}^2 dV_{G_{0,a}}}{\int_{\mathbb{M}_0} F^2 dV_{G_{0,a}}}. \quad (35)$$

Further, denoting the eigenfunction corresponding to $\Lambda_{k,a}^{\text{spat}}$ by F_k^{spat} let us denote $\mathbb{S}_k^{\text{spat}} = \{F \in \mathbb{S}_0^{\text{spat}} : \langle F, F_i^{\text{spat}} \rangle = 0, 1 \leq i \leq k\}$. We then have the variational characterisation of spatial eigenfunctions:

$$\Lambda_{k,a}^{\text{spat}} = - \inf_{F: \mathbb{M}_0 \rightarrow \mathbb{R}, F \in \mathbb{S}_{k-1}^{\text{spat}}} \frac{\int_{\mathbb{M}_0} \|\nabla_{G_{0,a}} F\|_{G_{0,a}}^2 dV_{G_{0,a}}}{\int_{\mathbb{M}_0} F^2 dV_{G_{0,a}}}. \quad (36)$$

⁵One has $0 = \int_{\mathbb{M}_0} F f \mathbf{1}_M = \int_0^\tau f(t) \int_M F(t, \cdot) d\ell dt$ for all $f \in W_0$, implying $\int_M F(\cdot, x) d\ell(x) \in W_0^\perp$.

Recall from (19) that the matrix representation of $G_{0,a}$ is

$$[G_{0,a}(t, x)] = \begin{pmatrix} 1/a^2 & 0 \\ 0 & g_t(x) \end{pmatrix}. \quad (37)$$

Thus, using (31) we have

$$\begin{aligned} \|\nabla_{G_{0,a}} F(t, x)\|_{G_{0,a}}^2 &= [\partial F(t, x)]^\top [G_{0,a}(t, x)]^{-1} [\partial F(t, x)] \\ &= a^2 (\partial_t F(t, x))^2 + [\partial_x F(t, x)]^\top [g_t(x)]^{-1} [\partial_x F(t, x)] \\ &= a^2 (\partial_t F(t, x))^2 + \|\nabla_{g_t} F(t, x)\|_{g_t}^2. \end{aligned} \quad (38)$$

Before stating our main result for this subsection, we note that because the spectrum of $\Delta_{G_{0,a}}$ can be written as the disjoint union $\sigma(\Delta_{G_{0,a}}) = \bigcup_{k \geq 1} \Lambda_{k,a}^{\text{temp}} \cup \bigcup_{k \geq 1} \Lambda_{k,a}^{\text{spat}}$, the indexing of the elements of $\sigma(\Delta_{G_{0,a}})$ immediately yields $\Lambda_{k,a}^{\text{temp}} \leq \Lambda_{k,a}$ and $\Lambda_{k,a}^{\text{spat}} \leq \Lambda_{k,a}$ for $k \geq 1$.

Theorem 7.

1. For each $k \geq 1$ and $a > 0$ one has $\lambda_k^D \leq \Lambda_{k,a}^{\text{spat}}$.
2. For each $k \geq 1$, $\Lambda_{k,a}$, $\Lambda_{k,a}^{\text{temp}}$, and $\Lambda_{k,a}^{\text{spat}}$ are nonincreasing in $a \geq 0$,
3. For each $k \geq 1$, $\lim_{a \rightarrow \infty} \Lambda_{k,a}^{\text{temp}} \rightarrow -\infty$.
4. For each $k \geq 1$, $\lim_{a \rightarrow \infty} \Lambda_{k,a}^{\text{spat}} = \lim_{a \rightarrow \infty} \Lambda_{k,a} = \lambda_k^D$.

Proof. See Appendix A. □

As a increases, part 2 of Theorem 7 states that $\Lambda_{k,a}^{\text{temp}}$ and $\Lambda_{k,a}^{\text{spat}}$ monotonically decrease. This is intuitive because a larger a leads to larger a value of $\|\nabla_{G_{0,a}} F\|_{G_{0,a}}^2$ in (35) as we increasingly penalise variation of F in the temporal direction. We note that as a increases, the ordering of eigenvalues $\Lambda_{k,a}$ in the full spectrum will change, and therefore the index k is implicitly a function of a . Temporal eigenvalues are demoted to lower positions in the full spectrum as a increases, leaving only spatial eigenvalues in the leading part of the full spectrum for sufficiently large a . Parts 1, 2, and 4 of Theorem 7 are illustrated numerically in Figure 3 for the Childress–Soward system from section 7.2.

For fixed k and increasing a , we expect the eigenfunction $F_{k,a}^{\text{spat}}$ to become more regular in the temporal direction as the infimum in (36) seeks to reduce the combination of gradients in the temporal and spatial directions. In the limit as $a \rightarrow \infty$, there will be vanishing variation in the temporal direction and we will recover the k^{th} eigenfunction of the dynamic Laplacian Δ^D , copied across time. In the other direction, as $a \rightarrow 0$, and the penalisation of the temporal variation diminishes, we expect $F_{k,a}^{\text{spat}}(t, \cdot)$ to simply encode the spatial structure of (M, g_t) ; that is, $F_{k,a}^{\text{spat}}(t, \cdot) \approx f_{k,t}$, where $f_{k,t}$ is the k^{th} eigenfunction of Δ_{g_t} .

From the above discussion we see that level sets of the eigenfunctions $F_{k,a}^{\text{spat}}$, from which we will create our semi-material finite-time coherent sets, will interpolate from being strictly material (in the $a \rightarrow \infty$ limit) to rather non-material (for small a). This is consistent with the discussion of the behaviour of stochastic trajectories at the conclusion of section 2.1, and the behaviour of the minimising disconnectors \mathbb{T} (which form boundaries of FTCS or part thereof) in Remark 6.

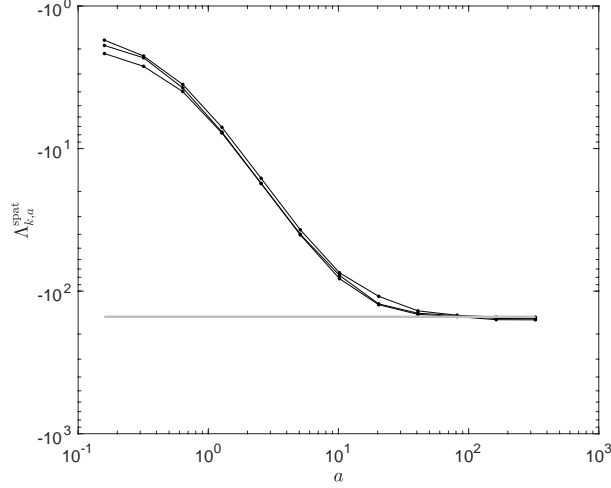


Figure 3.: The three subdominant spatial eigenvalues $\Lambda_2^{\text{spat}}, \Lambda_3^{\text{spat}}, \Lambda_4^{\text{spat}}$ of $\Delta_{G_{0,a}}$ versus a . The associated system is discussed in section 7.2. The grey horizontal lines (indistinguishable in this figure) indicate the values of the three subdominant eigenvalues of the dynamic Laplace operator Δ^D for this system.

4. Semi-material coherent sets from the inflated dynamic Laplacian

We shall now discuss how we can utilize the previous theoretical considerations to identify coherent sets and their lifetimes. In particular, we explain how we can identify different dynamical regimes—existence of coherent sets or global mixing—from the eigenmodes of the inflated dynamic Laplace operator $\Delta_{G_{0,a}}$. We illustrate these ideas using the partially coherent Childress–Soward system described in full detail in section 7.2.

4.1. Choosing the temporal diffusion parameter a

For large a , temporal diffusion in $\Delta_{G_{0,a}}$ will dominate and because of the variational (minimisation) characterisation of the eigenvalues, any temporal variation in the eigenfunctions will be heavily penalised. Therefore we expect eigenfunctions corresponding to eigenvalues early in the spectrum to be purely spatial. More precisely, from section 3.2 and Theorem 7, for large a we expect spatial eigenfunctions of $\Delta_{G_{0,a}}$ to be approximately “time-copied” versions of the eigenfunctions of the dynamic Laplace operator. In the other direction, for small a there is very low temporal diffusion and different time fibres of eigenfunctions $F(t, \cdot)$ will approximately decouple and depend almost entirely on the spatial metric g_t on the t^{th} time fibre. If one were to attempt to extract coherent sets through level sets of F in this small a regime, the coherent sets could be highly non-material.

We aim for a sweet spot for a somewhere in between these extremes. We would like to have the dominant eigenfunctions of $\Delta_{G_{0,a}}$ consisting mostly of spatial eigenfunctions, because it is these we are primarily interested in, but also including a small number of temporal eigenfunctions, so that such an a allows some temporal variation in the spatial

eigenfunctions. The latter point is crucial for being able to discriminate between coherent and mixing regimes over our full time domain.

We now discuss a heuristic to select a lower bound for a . The largest nonzero eigenvalue from the purely temporal component of $\Delta_{G_{0,a}}$ is $\Lambda_1^{\text{temp}} = -a^2\pi^2/\tau^2$, where τ is the flow duration. Assuming a rectangular domain M with (maximal) side length l , the largest nontrivial eigenvalue of the Laplace–Beltrami operator on (M, e) is $-4\pi^2/l^2$ for periodic boundary conditions, and $-\pi^2/l^2$ for homogeneous Neumann boundary conditions, respectively. The spatial eigenvalues of $\Delta_{G_{0,a}}$ will in general be larger in magnitude (more negative) than these values because of the presence of dynamics. Thus, if we desire the contribution from the temporal component to be about the same as the spatial component (with no dynamics), in the periodic case we want $a^2\pi^2/\tau^2 \approx 4\pi^2/l^2$, so we set $a_{\min} = 2\tau/l$ as the lower bound for a . Similarly, for homogeneous Neumann boundary conditions, we set $a_{\min} = \tau/l$. In section 7.2, this leads to $a_{\min} = 2/\pi$ for our numerical example. We suggest beginning with a_{\min} computed in this way and then increasing a_{\min} . Using this heuristic for the partially coherent Childress–Soward system in Section 7, one obtains a spectrum as shown in Figure 4. As predicted, we see that $\Lambda_2^{\text{spat}} < \Lambda_1^{\text{temp}}$

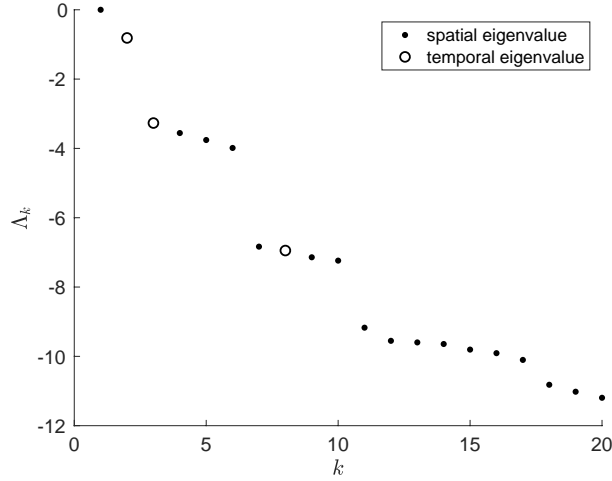


Figure 4.: Eigenvalues of $\Delta_{G_{0,a}}$ for the partially coherent Childress–Soward system for $a = 2/\pi$, estimated by the FEM-based discretisation described in section 6. Spatial modes are indicated by dots, temporal ones by circles.

4.2. Distinguishing spatial and temporal eigenfunctions

From section 3.4.2 we know that spatial eigenfunctions F of $\Delta_{G_{0,a}}$ have time fibres $F(t, \cdot)$ with constant mean, cf. (34). The temporal eigenfunctions have the form $F_k(t, x) = C \cos(k\pi t/\tau)$, $k \geq 1$. We will numerically compute the variance of the means of the time fibres $F(t, \cdot)$; if this variance is close to zero, the eigenfunction is spatial. We now analytically determine the variance of the temporal eigenfunctions F_k , $k \geq 1$. Let us normalise so that $\|F_k\|_{L^2(\mathbb{M}_0)}^2 = \tau\ell(M) = \|\mathbf{1}\|_{L^2(\mathbb{M}_0)}^2$; this implies $C = \sqrt{2}$. The mean $s(t)$

of $F(t, \cdot)$ on the t^{th} time fibre is

$$s(t) := \frac{1}{\ell(M)} \int_M C \cos(k\pi t/\tau) d\ell = \sqrt{2} \cos(k\pi t/\tau).$$

Toward computing the temporal variance of the spatial means $s(t)$, we note the mean of $s(t)$ is $\frac{1}{\tau} \int_0^\tau s(t) dt = 0$. Thus, the variance of $s(t)$ is

$$\frac{1}{\tau} \int_0^\tau (\sqrt{2} \cos(k\pi t/\tau))^2 dt = \frac{1}{\tau} \cdot 2 \frac{\tau}{2} = 1.$$

Therefore, with the above normalisation we have a simple numerical procedure for distinguishing spatial from temporal eigenfunctions by computing the variance of $s(t)$. If the variance of $s(t)$ is zero (or near zero), the eigenfunction is spatial and if the variance of $s(t)$ is 1 (or near 1), the eigenfunction is temporal. This scheme was used to categorise the spectrum shown in Figure 4.

4.3. Distinguishing coherent flow regimes from mixing regimes

The (signed) mass of a spatial mode F has to distribute itself over its time fibres $F(t, \cdot)$ because

$$\|F\|_{L^2(\mathbb{M}_0)}^2 = \int_0^\tau \|F(t, \cdot)\|_{L^2(M)}^2 dt = 1.$$

If the temporal diffusion coefficient a is suitably chosen, we will be able to distinguish temporal regions where coherent dynamics is present or absent using the L^2 norms of time fibres of *subdominant* eigenfunctions $F(t, \cdot)$, $t \in [0, \tau]$. It is important to recall that $\int_M F(t, \cdot) d\ell = 0$ for a.e. $t \in [0, \tau]$, which we have by (34) and the fact that subdominant spatial modes are also orthogonal to $\mathbf{1}_{\mathbb{M}_0}$. This implies that the only way for $F(t, \cdot)$ to be constant on the t^{th} time fibre is $F(t, \cdot) \equiv 0_{\{t\} \times M}$.

For t in intervals where coherent dynamics is present, the norm of $F(t, \cdot)$ may be relatively large, with $F(t, \cdot)$ taking large positive (say) values in the coherent region in space and negative values in the complement of the coherent region. Within each coherent region, $F(t, \cdot)$ should be approximately constant to achieve small values of $\|\nabla_{g_t} F(t, \cdot)\|_{L^2(M)}$. On the other hand, during periods of intense global mixing in space, for sufficiently large a it is likely that $\|F(t, \cdot)\|_{L^2(M)}^2$ will be small. This is because the metric g_t is rapidly varying in time and in order to achieve a minimal eigenvalue in the variational characterisation of eigenvalues (i.e. low values of $\|\nabla_{g_t} F(t, \cdot)\|_{L^2(M)}$) the eigenfunction F should also be rapidly varying in time to adapt to g_t . In opposition to this effect, if a is large enough, rapid variation of F in time will be costly in the temporal direction (i.e. large values of $|\partial_t F|$). The way out is for $F(t, \cdot)$ to be constant (i.e. zero) when strong globally mixing is present. This pushes the (signed) mass of F onto the most coherent time fibres and minimises the L^2 norm on strongly mixing time fibres; see Figure 5. Of course, the above analysis is strictly for spatial eigenfunctions F because the L^2 norms of time-fibres of temporal eigenfunctions vary dramatically in time.

In summary, as a basic indicator to discriminate between coherent vs mixing regimes we use the relative values of the L^2 norms of the time fibres of dominant spatial eigenfunctions.

This intuition is further formalized in the next section.

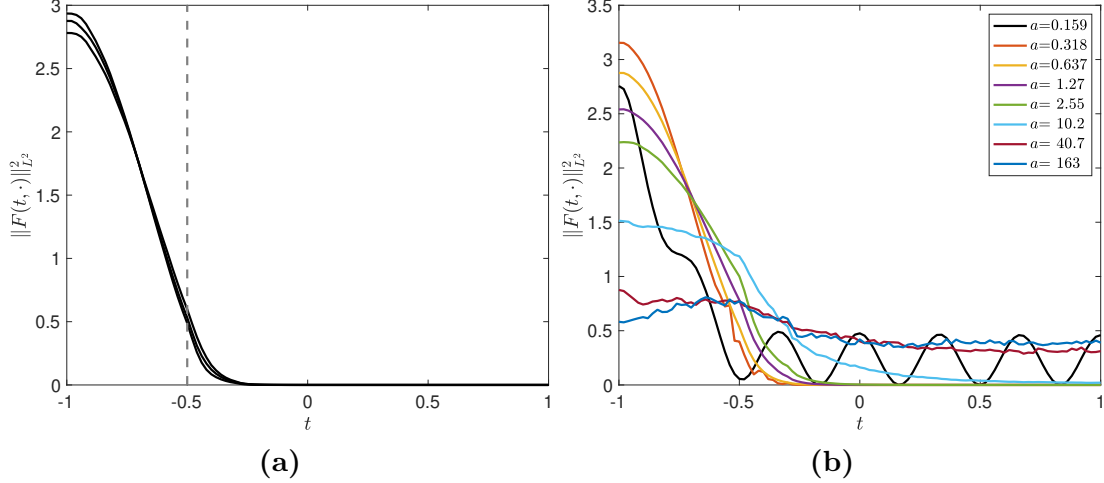


Figure 5.: Slicewise squared L^2 norms of subdominant spatial eigenfunctions for the partially coherent Childress–Soward system, which has a coherent regime for $t \in [-1, -0.5]$ and is mixing for $t \in [-0.5, 1]$. (a) The function $t \mapsto \|F(t, \cdot)\|_{L^2}^2$ for the first three subdominant spatial modes F for $a = \frac{2}{\pi}$, having the eigenvalues $\Lambda_k^{\text{spat}} = -3.5517, -3.7559, -3.9847$, $k = 2, 3, 4$. (b) Slicewise squared L^2 norms of the 4th spatial eigenmode Λ_4^{spat} of $\Delta_{G_{0,a}}$ for several different choices of the temporal diffusion strength a . We observe that the best distinction between the coherent and mixing regimes is obtained between $a = \frac{1}{\pi} \approx 0.32$ and $a = \frac{4}{\pi} \approx 1.27$ (recall our heuristic from subsection 4.1 suggested increasing a from the value $a_{\min} = \frac{2}{\pi}$). For $a \gtrsim \frac{8}{\pi}$, the temporal variation of F is too small and for $a \lesssim \frac{1}{2\pi}$ it is too large.

4.4. A posteriori considerations regarding a

The previous three subsections are sufficient to select a reasonable value for the parameter a , to separate temporal and spatial eigenfunctions, and to use the spatial eigenfunctions to find regimes of coherent behaviour. In this final subsection we take a closer look at the relationship between a and where in the spectrum a coherent set of a certain spatial regularity and temporal duration might be expected to appear. We will do this by comparing the dynamic Cheeger constant of a specified finite-time coherent set with the Cheeger constants of sets extracted from level sets of temporal eigenfunctions.

Consider a set $A \subset M$ that remains coherent in the time interval $0 < \tau_1 < \tau_2 < \tau$. By volume-preservation of the dynamics we write ℓ for the volume on M and later also for the volume on each time slice $\{t\} \times M$. Assuming that $\ell(A) \leq \ell(M)/2$ and following (24), but without taking the infimum, the dynamic Cheeger constant of the disconnecter $\Gamma = \partial A$ on the interval $[\tau_1, \tau_2]$ is:

$$h^D(\partial A) = \frac{\frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} V_{i^* g_t}(\partial A) dt}{\ell(A)},$$

where $i : A \hookrightarrow M$ is the inclusion map. The set A naturally defines a space-time set of the form $\mathbb{A} = [\tau_1, \tau_2] \times A \subset \mathbb{M}_0$. Because $\ell(A) \leq \ell(M)/2$, we have $V_{G_0,a}(\mathbb{A}) \leq V_{G_0,a}(\mathbb{M}_0)/2$, and thus the Cheeger constant of $\mathbb{I} = \partial\mathbb{A}$ is:

$$H_a(\partial\mathbb{A}) = \frac{2\ell(A) + \frac{1}{a} \int_{\tau_1}^{\tau_2} V_{i^*g_t}(\partial A) dt}{(\tau_2 - \tau_1)\ell(A)/a} = \frac{2a}{\tau_2 - \tau_1} + h^D(\partial A). \quad (39)$$

We now turn to the temporal eigenfunctions. We wish to construct a superlevel set from the k^{th} temporal eigenfunction $F_k^{\text{temp}}(t, x) = \cos(k\pi t/\tau)$ that has least Cheeger constant; this will occur for the level set at 0. We therefore define the superlevel set $\mathbb{A}_k^{\text{temp}} := \{(t, x) \in \mathbb{M}_0 : F_k^{\text{temp}}(t, x) > 0\}$. The boundary $\partial\mathbb{A}_k^{\text{temp}} = (F_k^{\text{temp}})^{-1}(0)$, consists of k copies of M . The associated spatiotemporal Cheeger constant is

$$H_a(\partial\mathbb{A}_k^{\text{temp}}) = \frac{k\ell(M)}{(\tau/2)\ell(M)/a} = \frac{2ka}{\tau}.$$

We now wish to compare $H_a(\partial\mathbb{A})$ with $H_a(\partial\mathbb{A}_k^{\text{temp}})$ and so we equate these two values. Solving the resulting equality for k yields

$$k = \tau \left(\frac{1}{\tau_2 - \tau_1} + \frac{h^D(\partial A)}{2a} \right). \quad (40)$$

We note a few points for fixed a .

- Coherent sets with shorter duration will tend to appear further down the spectrum because the term $\frac{\tau}{\tau_2 - \tau_1}$ increases with shrinking duration $\tau_2 - \tau_1$.
- A coherent set with a larger dynamic Cheeger constant on $[\tau_1, \tau_2]$ will appear further down the spectrum.

If we consider varying a :

- When a is small it is predominantly the dynamic Cheeger constant that governs where the set appears in the spectrum, and when a is larger, the temporal duration of the coherence is the important factor.
- Rearranging (40) to solve for a , we obtain

$$a = \frac{h^D(\partial A)}{2} \left(\frac{k}{\tau} - \frac{1}{\tau_2 - \tau_1} \right)^{-1}.$$

This provides a rough indication of a choice of a to pick up the coherent set A in a spatial eigenfunction appearing approximately nearby the k^{th} temporal eigenfunction in the eigenvalue ordering; note that smaller duration $\tau_2 - \tau_1$ will force larger k to maintain positivity of the second term above.

For the Childress–Soward flow from section 7.2 let us consider the second spatial mode, which identifies two vortices next to each other as a coherent set A ; see Figure 9(a). This set has perimeter 4π because the domain is periodic, and area $2\pi^2$; therefore $h^D(\partial A) =$

$2/\pi$. Further, $\tau = 2$ and $\tau_2 - \tau_1 = 1/2$, and we note that the boundary of $\partial\mathbb{A}$ at $t = -1$ does not enter the Cheeger constant calculations because of the Neumann boundary conditions on the temporal faces of \mathbb{M}_0 . Thus, $H_a(\mathbb{T}) = \frac{a}{T} + h^D(\partial A)$ in (39) and in (40) we replace $\tau_2 - \tau_1$ by $2(\tau_2 - \tau_1)$. The equality (40) becomes $k = 2 + \frac{2}{a\pi}$. This can be satisfied for $k \geq 3$, and with $k = 3$ it yields $a = 2/\pi$. This is the same value the heuristic from section 4.1 suggested.

5. A one-dimensional surrogate model

In this section we construct a reduced one-dimensional eigenproblem from the inflated dynamic Laplace eigenproblem by integrating out the spatial dynamics. The analysis of this reduced problem further formalizes our intuition from the previous section on how to deduce regimes of coherence and mixing.

5.1. Derivation of a surrogate 1D model

Let $\Delta_{G_{0,a}} F = \Lambda F$ with $\|F\|_{L^2(\mathbb{M}_0)} = 1$. Using (7), we multiply both sides by F and integrate over the t^{th} fibre $\{t\} \times M$:

$$\begin{aligned} \Lambda \|F(t, \cdot)\|_{L^2(M)}^2 &= \Lambda \int_M F^2 d\ell = a^2 \int_M (\partial_{tt} F) \cdot F d\ell + \int_M \Delta_{g_t} F \cdot F d\ell \\ &= a^2 \int_M \frac{1}{2} \partial_{tt}(F^2) - (\partial_t F)^2 d\ell - \int_M \|\nabla_{g_t} F\|_{g_t}^2 d\ell \\ &= \frac{a^2}{2} \partial_{tt} \int_M F^2 d\ell - a^2 \int_M (\partial_t F)^2 d\ell - \int_M \|\nabla_{g_t} F\|_{g_t}^2 d\ell \end{aligned}$$

which we write, using $u(t) = \|F(t, \cdot)\|_{L^2(M)}^2$, as

$$\frac{a^2}{2} u'' - \left(a^2 \frac{\int_M (\partial_t F)^2 d\ell}{\|F(t, \cdot)\|_{L^2(M)}^2} + \frac{\int_M \|\nabla_{g_t} F\|_{g_t}^2 d\ell}{\|F(t, \cdot)\|_{L^2(M)}^2} \right) u = \Lambda u.$$

Making the obvious substitutions for the (time-dependent) temporal and spatial Rayleigh-type coefficients, we write this as

$$\frac{a^2}{2} u''(t) = (\Lambda + [a^2 \rho^{\text{temp}}(t) + \rho^{\text{spat}}(t)]) u(t), \quad \text{for } t \in (0, \tau), \quad u'(0) = u'(\tau) = 0. \quad (41)$$

Our reduced equation (41) describes the expected behaviour of $u(t) = \int_M F(t, \cdot)^2 d\ell$, the square of the spatial norm of the eigenfunction F on the t^{th} time slice. On the t^{th} time fibre, the decay experienced due to the irregularity of F is $a^2 \rho^{\text{temp}}(t) + \rho^{\text{spat}}(t)$. We interpret Λ as the average space-time decay that the eigenfunction F experiences on all of \mathbb{M}_0 . Indeed, by the variational form (35) we have that

$$-\Lambda = \int_0^\tau (a^2 \rho^{\text{temp}}(t) + \rho^{\text{spat}}(t)) u(t) dt,$$

and $\int_0^\tau u(t) dt = 1$ by our choice of normalization of F .

Recall that in section 4.3 we used the relative size of the fibre norms $u(t)$ to distinguish coherent flow regimes from incoherent ones. There are two fundamental regimes:

1. $t \in [0, \tau]$ for which $\Lambda + (a^2 \rho^{\text{temp}}(t) + \rho^{\text{spat}}(t)) < 0$. For such t the *local decay is less than the average decay*, indicative of F encoding *relatively coherent dynamics*. Because $u'' < 0$ and $u > 0$, u has a local maximum. In other words there is a local peak in the norm of $\|F(t, \cdot)\|_{L^2(M)}$, consistent with the discussion in section 4.3.
2. $t \in [0, \tau]$ for which $\Lambda + (a^2 \rho^{\text{temp}}(t) + \rho^{\text{spat}}(t)) > 0$. For such t the *local decay is greater than the average decay*, indicative of F encoding *relatively mixing dynamics*. Because $u'' > 0$ and $u > 0$, u (and therefore $\|F(t, \cdot)\|_{L^2(M)}$) has a local minimum, consistent with the discussion in section 4.3.

The above two regimes partition $[0, \tau]$ into time intervals where the eigenfunction F encodes dynamics that is more coherent or less coherent, respectively, than the average coherence over all of $[0, \tau]$. One could also define subintervals of τ with more extreme coherence relative to F by introducing a threshold $c > 0$. For example, the sets $\{t \in [0, \tau] : \Lambda + (a^2 \rho^{\text{temp}}(t) + \rho^{\text{spat}}(t)) < -c\}$ and $\{t \in [0, \tau] : \Lambda + (a^2 \rho^{\text{temp}}(t) + \rho^{\text{spat}}(t)) > c\}$ indicate stronger coherence and stronger mixing, respectively, with increasing c . On the former interval, $\|F(t, \cdot)\|_{L^2(M)}$ has a local maximum and on the latter, $\|F(t, \cdot)\|_{L^2(M)}$ has a local minimum. In the next subsection we analyse the shape of, and transitions between, these maxima and minima.

5.2. Analysis of a surrogate 1D model

In the previous subsection, the coefficient function $a^2 \rho^{\text{temp}}(t) + \rho^{\text{spat}}(t)$ arose directly from the eigenfunction F . We now heuristically investigate replacing this exact coefficient function with a function denoted simply $\rho(t)$, whose form is suggested by properties of the flow, in an effort to infer something about F . To this end, we write a schematic version of (41),

$$\frac{a^2}{2} u''(t) - \rho(t) u(t) = \nu u(t), \quad \text{for } t \in (0, \tau), \quad u'(0) = u'(\tau) = 0, \quad (42)$$

where $\rho(t) \geq 0$ is meant to describe the “relative mixing strength” (larger ρ , greater mixing) that the flow inflicts on the supposed unknown function F on the t^{th} time fibre.

We assume that all we know in (42) is a and ρ , and so this equation amounts to a *Sturm–Liouville eigenproblem*. By the theory of Sturm–Liouville eigenproblems [Tes12, §5.3–§5.4, pp. 153 and 164, and Thm. 5.17], if ρ is integrable, (42) has a countable spectrum of distinct eigenvalues $0 \geq \nu_0 > \nu_1 > \dots$ all having multiplicity one, and the associated (up to constant scaling unique) eigenfunctions u_i are mutually orthogonal in $L^2([0, \tau])$ and have exactly i zeros, $i \geq 0$. Since u models the squared norm of time slices of F , only solutions $u_i \geq 0$ are of interest, which leaves u_0 as the unique meaningful solution. We note that if we have an eigenfunction F , applying the above remarks to (41), which we obtain by substituting $\rho = a^2 \rho^{\text{temp}} + \rho^{\text{spat}}$ and $\nu = \Lambda$ into (42), shows that the solution $u(t) = \|F(t, \cdot)\|_{L^2(M)}^2$ is the unique solution.

Returning to our heuristic discussion, for the Childress–Soward flow, from the discussion in section 4.3 and Figure 5, we expect $\rho(t)$ arising from eigenfunctions F that highlight the coherent sets from time -1 to -0.5 to be small until the mixing regime begins at $t = -0.5$, after which $\rho(t)$ should rise to a much larger value. A simple approximation of such a ρ is a step function with two values $Z \gg z > 0$ in the coherent and mixing regimes, respectively. This step function form of $\rho(t)$ permits finer analysis of the surrogate model (42).

For simplicity, in the following we set $\tau = 1$. It is straightforward to compute a one-to-one correspondence between the following two homogeneous Neumann boundary value problems, one on $[0, \tau]$, and one on $[0, 1]$:

$$\frac{\tilde{a}^2}{2} \tilde{u}'' - \tilde{\rho} \tilde{u} = \nu \tilde{u} \text{ on } (0, \tau) \quad \xleftrightarrow[\tilde{a} = a\tau, \tilde{\rho}(t) := \rho(t/\tau)]{\tilde{u}(t) := u(t/\tau)} \quad \frac{a^2}{2} u'' - \rho u = \nu u \text{ on } (0, 1) \quad (43)$$

The solution of the problem on $[0, \tau]$ is obtained from the solution on $[0, 1]$ with scaled variables, if the temporal diffusion strength is also scaled by τ .

As in the Childress–Soward flow in Section 7 we assume that the velocity field of the system is such that there is coherence in the first $0 < p < 1$ fraction of the time interval, and then there is strong mixing. Of course, the coherent regime could be located wherever in the time interval, our choice is merely for simplicity. We set

$$\rho(t) = \begin{cases} z, & t \in [0, p], \\ Z, & t \in (p, 1], \end{cases} \quad (44)$$

with $Z \gg z > 0$. We recognise that replacing the a priori unknown coefficient function ρ in (41) by a ρ taking only two values is a strong simplification. The solutions of (42) can now be determined analytically, and the numerical results in Figure 6(a) and Figure 7 show that the profile of u predicted by the surrogate model with this idealised ρ is surprisingly accurate.

Proposition 8. *The solutions to (42) with mixing rate function ρ as in (44) are*

$$u(t) = \begin{cases} \alpha \cosh(\omega_z t), & t \in [0, p], \\ \alpha \frac{\cosh(\omega_z p)}{\cosh(\omega_Z(1-p))} \cosh(\omega_Z(1-t)), & t \in [p, 1], \end{cases}$$

where α is an arbitrary scaling constant, $\omega_z = \omega_z(\nu) = \frac{\sqrt{2(\nu+z)}}{a}$, $\omega_Z = \omega_Z(\nu) = \frac{\sqrt{2(\nu+Z)}}{a}$, and ν is the eigenvalue satisfying

$$f(\nu) := \frac{\omega_z(\nu)}{\omega_Z(\nu)} \tanh(\omega_z(\nu) p) + \tanh(\omega_Z(\nu) (1-p)) = 0. \quad (45)$$

Proof. See Appendix B.1. □

Recall from above that we are only interested in the dominant mode of (42), i.e., the eigenfunction associated with ν_0 , the largest eigenvalue. Properties of the eigenproblem are discussed in Appendix B.2. In particular, we show that $\nu_0 \in (-Z, -z)$, and this

implies $\omega_z(\nu_0) \in i\mathbb{R}$ and $\omega_Z(\nu_0) \in \mathbb{R}$. Consequently, for the associated eigenfunction u we obtain a cosine-profile on the first (coherent) part $[0, p]$ of the time interval because $\cosh(\omega_z t) = \cos(|\omega_z|t)$, and an exponentially decaying cosh-profile on the second (mixing) part $[p, 1]$.

The differences in magnitude of u will, ideally, be indicative for the difference between the regimes. In particular, we expect strong exponential decay of u in a strongly mixing regime. Nevertheless, ambiguity for the intermediate times when the system is shifting from coherent to mixing, is still expected, as for most (realistic) systems this is a continuous and not an abrupt transition.

We show the dominant eigenfunction u from Proposition 8 for parameters $a^2/2 \in \{1/\pi^2, 1/100\pi^2\}$, $p = 0.25$, $z = 2$, $Z = 40$ in Figure 6(a).

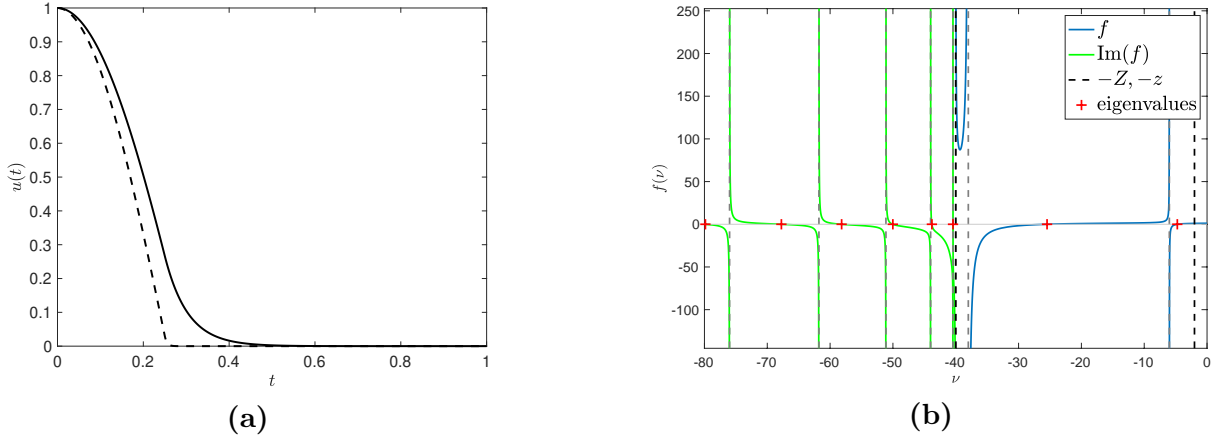


Figure 6.: (a) Example solution of (42), as in Proposition 8 below, for parameters $a^2/2 = 1/\pi^2$, $p = 0.25$, $z = 2$, $Z = 40$ (solid). This matches well with the full solutions shown in Figure 5; note that the time axis has been linearly rescaled from $[-1, 1]$ to $[0, 1]$ in this figure (see (43)). Decreasing a to $a^2/2 = 1/100\pi^2$ (dashed) leads to a sharper transition in the surrogate solution. (b) Eigenvalues of (42) as zeros of (45) for $a^2/2 = 1/\pi^2$. The black dashed lines indicate the values $-Z = -40$, $-z = -2$. Gray dashed lines indicate singularities of f .

We find that the decay of u is increasingly rapid on $[p, 1]$ as a decreases. Note that this analysis assumes that ρ is independent of a in (41).

The surrogate model gives a good qualitative approximation, as shown in Figure 7. Subfigure (a) tests the cosine-profile by looking at whether $\arcsin(\|F(t, \cdot)\|_{L^2}^2)$ vs t is indeed linear for the coherent regime. We use arcsin instead of arccos to map near zero values to near zero values. Subfigure (b) tests approximately exponential decay of $\|F(t, \cdot)\|_{L^2}^2$ vs t in the mixing regime.

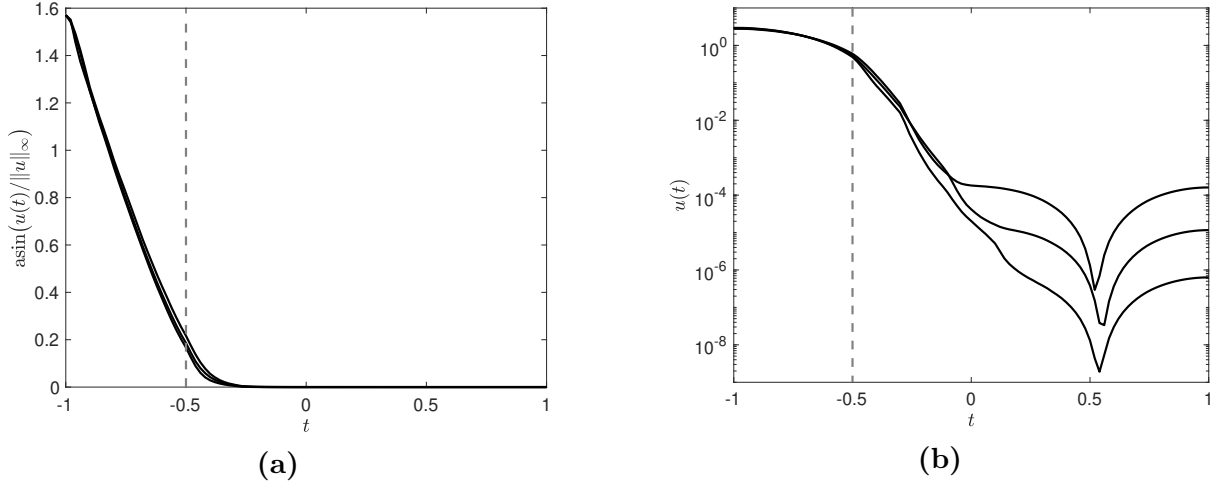


Figure 7.: Analysis of the function $u : t \mapsto \|F(t, \cdot)\|_{L^2(M)}^2$, for the first three spatial eigenmodes F_k^{spat} , $k = 2, 3, 4$, of $\Delta_{G_{0,a}}$ for the partially coherent Childress–Soward system with $a = \frac{2}{\pi}$. (a) $t \mapsto \arcsin(u(t)/\|u\|_\infty)$. Dashed lines indicate the time $t = -0.5$, where the mixing regime starts. The approximately constant slope on $[-1, -0.5]$ indicates that u is approximately a cosine on this interval where the velocity field is in the coherent regime. (b) The function u depicted on a semilogarithmic scale. The strong exponential drop in the graph over $[-0.5, -1]$ indicates the onset of mixing. Compare also with Figure 5.

6. FEM-based numerical discretisation of the inflated dynamic Laplacian

In this section we describe how to numerically approximate $\Delta_{G_{0,a}}$ in (7). We adopt and extend the approach of [FJ18], which derived a finite-element method discretisation of the dynamic Laplacian eigenproblem. For $f : M \rightarrow \mathbb{R}$ we define the pushforward and pullback of f under ϕ_t by $(\phi_t)_* f := f \circ (\phi_t)^{-1}$ and $\phi_t^* f := f \circ \phi_t$, respectively; recall that $g_t = \phi_t^* e$. The right-hand side of (7) has two main components and we begin by discussing Δ_{g_t} . We note that one has the alternative representation [Fro15], $\Delta_{g_t} = \phi_t^* \circ \Delta_{(\phi_t(M), e)} \circ (\phi_t)_*$, where we have made explicit the fact that the Laplace operator on the right-hand side is acting on the “future” Riemannian manifold $(\phi_t(M), e)$. Similarly for $F \in H^1(\mathbb{M}_0)$ (resp. $F \in H^1(\mathbb{M}_1)$) we have $\Phi_* F = F \circ \Phi^{-1} \in H^1(\mathbb{M}_1)$ (resp. $\Phi^* F = F \circ \Phi \in H^1(\mathbb{M}_0)$). We are now ready to construct a weak-form approximation of the eigenproblem $\Delta_{G_{0,a}} F = \Lambda F$, where for the moment we assume homogeneous Neumann conditions on M (M may also be boundaryless). Neumann boundary conditions [Fro15, FJ18] on M allow us to find finite-time coherent sets that may share a boundary with M , while Dirichlet boundary conditions [FJ18] on M force the finite-time coherent sets to have boundaries away from the boundary of M . In the weak form, the only change required to solve the Dirichlet case is to change $H^1(\mathbb{M}_0)$ to $H_0^1(\mathbb{M}_0)$. For $F, \tilde{F} \in H^1(\mathbb{M}_0)$, multiplying our eigenproblem by \tilde{F} and integrating both sides we obtain

$$\int_0^\tau \int_M \Delta_{G_{0,a}} F \cdot \tilde{F} \, d\ell \, ds = \Lambda \int_0^\tau \int_M F \cdot \tilde{F} \, d\ell \, ds. \quad (46)$$

The left-hand side of (46) is

$$\begin{aligned}
& \int_0^\tau \int_M (a^2 \partial_{tt}|_{t=s} F(\cdot, x)) \tilde{F}(\cdot, x) \, d\ell(x) \, ds + \int_0^\tau \int_M (\Delta_{g_s} F(s, \cdot)) \tilde{F}(\cdot, s) \, d\ell \, ds \\
&= \int_0^\tau \int_M (a^2 \partial_{tt}|_{t=s} F(\cdot, x)) \tilde{F}(\cdot, x) \, d\ell(x) \, ds \\
&\quad + \int_0^\tau \int_{\phi_s(M)} (\Delta_{(\phi_s(M), e)} \circ (\Phi_* F)(s, \cdot)) \cdot (\Phi_* \tilde{F})(s, \cdot) \, d\ell \, ds \\
&= -a^2 \int_0^\tau \int_M (\partial_t|_{t=s} F(\cdot, x)) \cdot \left(\partial_t|_{t=s} \tilde{F}(\cdot, x) \right) \, d\ell(x) \, ds \tag{47}
\end{aligned}$$

$$- \int_0^\tau \int_{\phi_s(M)} \nabla_x(\Phi_* F) \cdot \nabla_x(\Phi_* \tilde{F}) \, d\ell \, ds. \tag{48}$$

Similarly, the right-hand side of (46) is

$$\Lambda \int_0^\tau \int_{\phi_s(M)} \Phi_* F \cdot \Phi_* \tilde{F} \, d\ell \, ds. \tag{49}$$

Because of the differing roles of time and space we assume that our approximating basis $V \subset H^1([0, \tau] \times M)$ contains functions of the form $\xi(t)\eta(x)$, $t \in [0, \tau], x \in M$. This ansatz enables a convenient decomposition across time, where at each time fibre we can leverage the spatial constructions from [FJ18]. It also allows for simple adjustment of the parameter a , without having to recompute any integrals. We suppose our approximation space is built in this way using a finite number of basis elements ξ_i , $i = 0, \dots, T$ and η_k , $k = 1, \dots, N$. Inserting $F(t, x) = \xi_i(t)\eta_k(x)$, $\tilde{F}(t, x) = \xi_j(t)\eta_l(x)$ into (47) and (48), we have

$$\begin{aligned}
(47) + (48) &= -a^2 \int_0^\tau \int_M \eta_k \partial_t|_{t=s} \xi_i \cdot \eta_l \partial_t|_{t=s} \xi_j \, d\ell \, ds \\
&\quad - \int_0^\tau \int_{\phi_s(M)} \xi_i(s) \nabla_x((\phi_s)_* \eta_k) \cdot \xi_j(s) \nabla_x((\phi_s)_* \eta_l) \, d\ell \, ds \tag{50}
\end{aligned}$$

$$\begin{aligned}
&= a^2 \int_0^\tau -\xi'_i(s) \cdot \xi'_j(s) \underbrace{\int_{\phi_s(M)} (\phi_s)_* \eta_k \cdot (\phi_s)_* \eta_l \, d\ell}_{=: M_{kl}^s} \, ds \tag{51}
\end{aligned}$$

$$+ \int_0^\tau \xi_i(s) \xi_j(s) \underbrace{\int_{\phi_s(M)} -\nabla_x((\phi_s)_* \eta_k) \cdot \nabla_x((\phi_s)_* \eta_l) \, d\ell}_{=: D_{kl}^s} \, ds =: \mathbf{D}_{ij,kl} \tag{52}$$

Similarly, we have

$$\frac{1}{\Lambda} (49) = \int_0^\tau \xi_i(s) \xi_j(s) \underbrace{\int_{\phi_s(M)} ((\phi_s)_* \eta_k) \cdot ((\phi_s)_* \eta_l) \, d\ell}_{=: M_{kl}^s} \, ds =: \mathbf{M}_{ij,kl} \tag{53}$$

We now fix the ξ_i , $i = 0, \dots, T$ and η_k , $k = 1, \dots, N$ to be the standard piecewise linear one-dimensional and d -dimensional hat functions, respectively. More precisely, we

partition $[0, \tau]$ into intervals with endpoints $0 = t_0 < t_1 < \dots < t_T = \tau$ and let ξ_i be the nodal hat function centred at node t_i . Given vertices $x_k \in M$, we mesh M into simplices and define η_k as the nodal hat function with node x_k ; this mesh is used to create $M_{kl}^{t_0}$ and $D_{kl}^{t_0}$. To define M^{t_i} and D^{t_i} , $i = 0, \dots, T$ we refer the reader to [FJ18]. For each $i = 0, \dots, T$, the matrices $M_{kl}^{t_i}$ are the standard mass matrices from the finite-element method, and the $D_{kl}^{t_i}$ are the modified stiffness matrices discussed in [FJ18], and can be efficiently computed.

To estimate $\mathbf{D}_{ij,kl}$ and $\mathbf{M}_{ij,kl}$, for $s \in [t_i, t_{i+1}]$ we linearly interpolate to estimate $M^s \approx M^{t_i} + ((s - t_i)/(t_{i+1} - t_i))(M^{t_{i+1}} - M^{t_i})$, similarly $D^s \approx D^{t_i} + ((s - t_i)/(t_{i+1} - t_i))(D^{t_{i+1}} - D^{t_i})$. Using these estimates, all that remains is to analytically evaluate the one-dimensional integrals (51)–(52) and (53). We omit the elementary, but lengthy details and present here the resulting formulae, where we specialise to $t_i = i\tau/T$ and set $h = \tau/T$. Note that because the one-dimensional functions u_i only overlap when $i = j$ or $|i - j| = 1$, it is only for these combinations of i and j that we obtain nonzero integrals.

$$\mathbf{D}_{ij,kl} = \begin{cases} -\frac{a^2}{2h}(M_{kl}^{t_0} + M_{kl}^{t_1}) + \frac{h}{12}(3D_{kl}^{t_0} + D_{kl}^{t_{i+1}}), & i = j = 0; \\ -\frac{a^2}{2h}(M_{kl}^{t_{i-1}} + 2M_{kl}^{t_i} + M_{kl}^{t_{i+1}}) \\ \quad + \frac{h}{12}(D_{kl}^{t_{i-1}} + 6D_{kl}^{t_i} + D_{kl}^{t_{i+1}}), & 1 \leq i = j \leq T - 1; \\ -\frac{a^2}{2h}(M_{kl}^{t_T} + M_{kl}^{t_{T-1}}) + \frac{h}{12}(3D_{kl}^{t_T} + D_{kl}^{t_{T-1}}), & i = j = T; \\ \frac{a^2}{2h}(M_{kl}^{t_i} + M_{kl}^{t_{i+1}}) + \frac{h}{12}(D_{kl}^{t_i} + D_{kl}^{t_{i+1}}), & j = i + 1, i \leq T - 1. \end{cases} \quad (54)$$

and

$$\mathbf{M}_{ij,kl} = \begin{cases} \frac{h}{12}(3M_{kl}^{t_0} + M_{kl}^{t_{i+1}}), & i = j = 0; \\ \frac{h}{12}(M_{kl}^{t_{i-1}} + 6M_{kl}^{t_i} + M_{kl}^{t_{i+1}}), & 1 \leq i = j \leq T - 1; \\ \frac{h}{12}(3M_{kl}^{t_T} + M_{kl}^{t_{T-1}}), & i = j = T; \\ \frac{h}{12}(M_{kl}^{t_i} + M_{kl}^{t_{i+1}}), & j = i + 1, i \leq T - 1. \end{cases} \quad (55)$$

The values for $j = i - 1, i = 1, \dots, T$ are identical to the values for $j = i + 1$ by symmetry. Thus, numerically we solve the sparse, symmetric eigenproblem $\mathbf{D}\mathbf{w} = \lambda\mathbf{M}\mathbf{w}$, where $\mathbf{w} \in \mathbb{R}^{(T+1)N}$. An approximate eigenfunction F is then reconstructed as $F(t, x) = \sum_{i=0}^T \sum_{k=1}^N \mathbf{w}_{i,k} \xi_i(t) \eta_k(x)$.

7. Numerical example

7.1. The Childress–Soward “cat’s-eye” flow

We consider the two-dimensional velocity field [CS89] $v : \mathbb{T}^2 \rightarrow \mathbb{R}^2$, parameterised by $A > 0$ and $-1 \leq r \leq 1$:

$$v = A \cdot \left(\frac{\partial \psi}{\partial y}, -\frac{\partial \psi}{\partial x} \right), \text{ with streamfunction } \psi(x, y) = \sin x \sin y + r \cos x \cos y, \quad (56)$$

where \mathbb{T}^2 is identified with $2\pi S^1 \times 2\pi S^1$, and S^1 is the circle of circumference 1. For $|r| \approx 1$ the flow is a diagonal shear and for $r \approx 0$ the flow has four vortices; otherwise it

possesses an intermediate “cat’s-eye” structure. Figure 8 shows streamfunctions of the flow for different values of r .

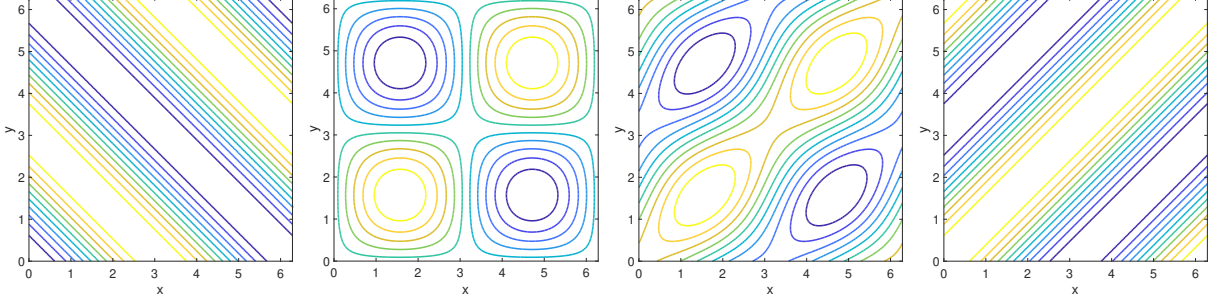


Figure 8.: Streamfunction contour lines of the Childress–Soward flow (56) for $r = -1, 0, 0.5, 1$, left to right. Light yellow colors indicate larger values. At any point, the velocity field v is tangential to contour lines of the streamfunction.

7.2. Extracting semi-material coherent sets from a flow exhibiting both coherent and incoherent regimes

We consider a non-autonomous Childress–Soward flow (56) with time-dependent amplitude $A(t) = 40 \mathbf{1}_{[-1,0.5]}(t) + 30 \mathbf{1}_{(-0.5,1]}(t)$ and time-dependent parameter modulation

$$r(t) = \begin{cases} 0, & t \in [-1, -0.5], \\ \text{sign}(\cos(5\pi t)), & t \in (-0.5, 1]. \end{cases}$$

This flow shows coherent behavior in the four vortices throughout the time interval $[-1, -0.5]$, while four alternating (perpendicularly) shearing periods on $[-0.5, 1]$, each of length approximately 0.4, result in a mixing regime on this latter time interval.

Standard LCS methods of coherent structure detection will *fail to detect the coherent behaviour* because it only lasts for the first quarter of the time duration $[-1, 1]$, and is then destroyed. We will show that *we can detect the coherent regime and corresponding coherent sets using the spectrum and eigenfunctions of $\Delta_{G_{0,a}}$* .

We generate trajectories of the system on the time interval $[-1, 1]$ sampled on a uniform grid of 101 time instances, with a 35×35 regular spatial grid of initial conditions. The approximation of $\Delta_{G_{0,a}}$ is carried out by the FEM-based method described in section 6. We select $a = 2/\pi$, based on the heuristic in section 4.1, and compute the leading 20 eigenvalues and corresponding eigenfunctions. Next, we separate the temporal eigenfunctions from the spatial ones, as described in section 4.2. Figure 4 shows the eigenvalues together with their types (spatial/temporal). Our choice of a seems to be appropriate because the leading few eigenvalues contain a small number of temporal eigenvalues within several spatial eigenvalues, the latter being our main interest. From Figure 4 we immediately see that in addition to F_1 , which is always spatial, the next spatial eigenfunctions are F_4, F_5, F_6 , followed by a clear gap in the spectrum to the next spatial eigenfunction F_7 . Thus, we expect to see four dominating coherent sets of similar coherence strength due to the similarity of the values of $\Lambda_4, \Lambda_5, \Lambda_6$.

Following the discussion in section 4.3, the relative size of the L^2 norm between different timeslices of our three subdominant spatial eigenfunctions $F_k(t, \cdot)$, $k = 4, 5, 6$ gives our first indication of the time durations over which we have coherent dynamics. Figure 5(a) shows relatively large and approximately equal values for these norms for each $k = 4, 5, 6$, in a time interval approximately equal to $[-1, -0.6]$, indicating possible coherence during this time interval. A precise time interval for coherence is not clear, but we can confidently say (i) the flow contains highly coherent sets at $t = -1$, (ii) there are likely 4 highly coherent sets encoded in F_1, F_4, F_5, F_6 , and (iii) that the coherence of these sets is lost by $t = -0.5$. Figure 5(b) shows the slicewise L^2 norms $t \mapsto u(t)$ of the dominant dynamic eigenmode for the choices $a = 2^l/\pi$, $l = -1, 0, 1, 2, 3, 5, 7, 9$, illustrating that the best choices of a lie between $1/\pi$ and $4/\pi$, consistent with our heuristics in sections 4.1 and 4.2.

Figure 9 shows timeslices of the first three subdominant spatial eigenmodes F_k^{spat} , $k = 2, 3, 4$. In this example, these correspond to the indices $k = 4, 5, 6$ in the global order; i.e. $F_2^{\text{spat}} = F_4$, $F_3^{\text{spat}} = F_5$, and $F_4^{\text{spat}} = F_6$. We fix the colorscale from the first timeslice $F_k^{\text{spat}}(-1, \cdot)$ as the slice-wise norms maximize there and thus indicate coherence. The first column in Figure 9 displays the eigenvectors F_k^{spat} for $k = 2, 3, 4$, in the first three rows, respectively. An automated way to separate the different coherent sets encoded by groups of eigenvectors is implemented in the Sparse Eigenbasis Approximation (SEBA) algorithm [FRS19]. It computes a rotation of leading eigenvectors such that the resulting rotated vectors are maximally sparse. SEBA is applied to the first four spatial eigenfunctions $F_1^{\text{spat}}, \dots, F_4^{\text{spat}}$, recalling that $F_1^{\text{spat}} \equiv \mathbf{1}_{M_0}$ by definition, to produce S_1, \dots, S_4 , which are functions of t and x . Each of these SEBA functions should be supported on a *single* semi-material coherent set and the value represents the relative strength of membership in one of the four coherent sets. Having separated the semi-coherent sets, we define $S_{\max} := \max\{S_1, \dots, S_4\}$ into a single function via superposition. The final row of the first column in Figure 9 shows $S_{\max}(-1, \cdot)$.

Column 2 of Figure 9 displays $F_2^{\text{spat}}(-0.5, \cdot)$, $F_3^{\text{spat}}(-0.5, \cdot)$, $F_4^{\text{spat}}(-0.5, \cdot)$, and $S_{\max}(-0.5, \cdot)$, respectively. Notice that by time $t = -0.5$, the point at which the dynamics enters its mixing phase, the relative magnitude of the eigenfunctions and SEBA superposition has begun to decrease; compare this with Figure 5(a). The third column of Figure 9 is the same as the second column, except t has been advanced from -0.5 to -0.4 , after which the mixing regime begins. A significant decrease in the magnitude of the eigenfunctions and SEBA superposition can now be seen, in line with Figure 5(a).

The fourth column of Figure 9 shows the pushforwards $((\phi_{0.5})_* F_2^{\text{spat}}(-1, \cdot), (\phi_{0.5})_* F_3^{\text{spat}}(-1, \cdot), (\phi_{0.5})_* F_4^{\text{spat}}(-1, \cdot), (\phi_{0.5})_* S_{\max}(-1, \cdot))$, respectively. In other words we take the functions in column 1 of Figure 9, fix the colour of each point in M and flow every point forward for 0.5 time units without changing its colour. The fourth column of Figure 9 clearly shows coherence on the time interval $[-1, -0.5]$. Advancing this flow a little further by 0.1 time units we arrive at the fifth column of Figure 9. One sees a dramatic difference, with rapid destruction of the coherent sets. This is strongly consistent with the small slicewise norm values shown in the third column, indicative of a lack of coherence.

The eigenmodes S_k^{spat} , $k = 2, 3, 4$ separate the four gyres from one another; higher spa-

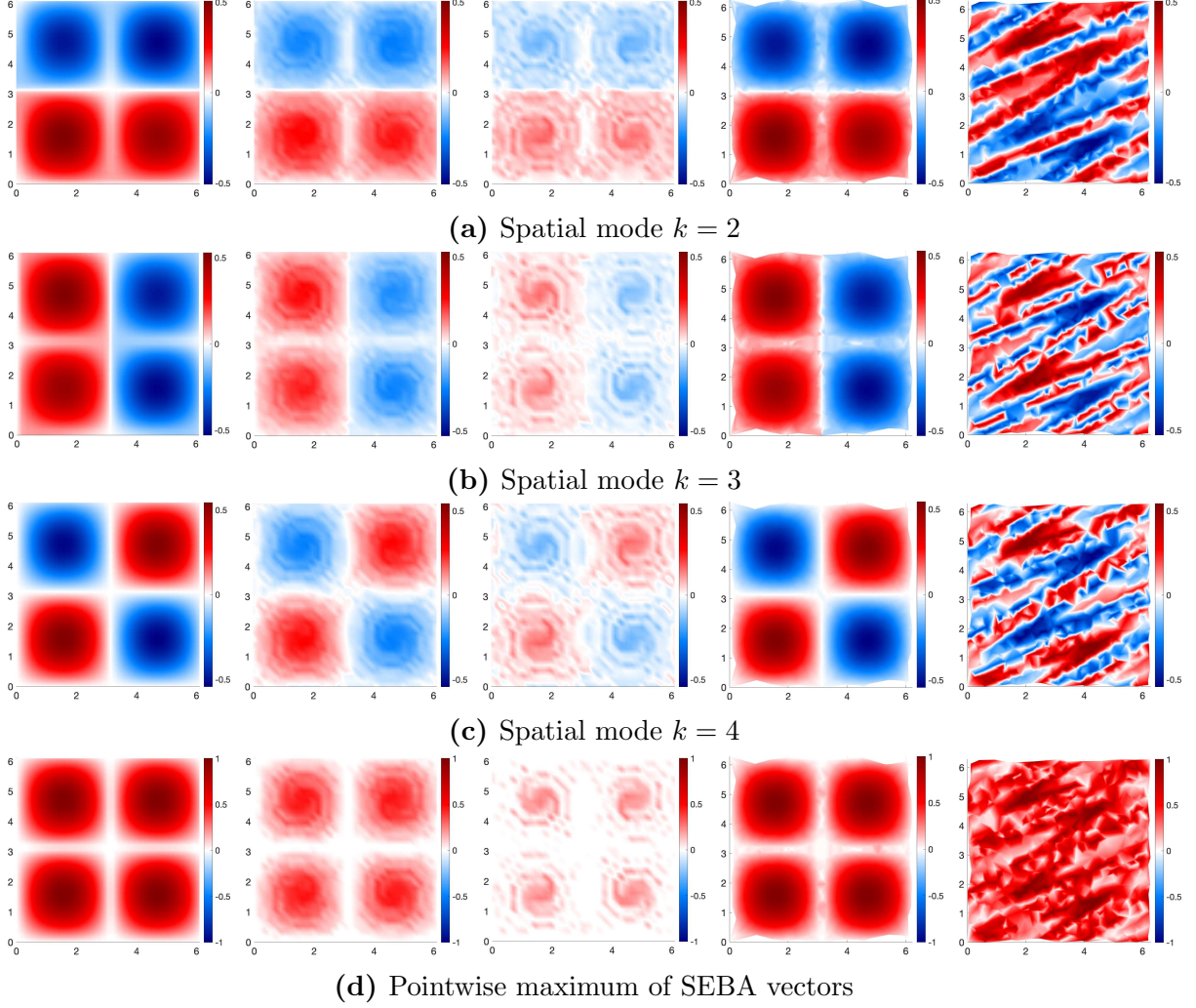


Figure 9.: Visualization of timeslices $F_k^{\text{spat}}(t, \cdot)$ of spatial eigenfunctions of $\Delta_{G_{0,a}}$, for $k = 2, 3, 4$ (rows (a), (b), and (c), respectively). Colours indicate the value of the time slice eigenfunction F_k^{spat} . The first three columns show slices of a function on the t^{th} time fibre of \mathbb{M}_0 for $t = -1, -0.5, -0.4$, respectively. The last two columns depict $F_k^{\text{spat}}(-1, \cdot)$ evolved forward to time fibres in \mathbb{M}_1 at $t = -0.5$ and -0.4 ; see the main text for a precise description. Row (d) shows the images corresponding to the previous rows for the SEBA superposition S_{max} .

tial modes subdivide the gyres into coherent rings and spiral-like structures (not shown). These are less coherent than the gyre structures in Figure 9, as indicated by the spectrum of $\Delta_{G_{0,a}}$. We summarise our approach in algorithm form below.

Algorithm (to extract semi-material coherent sets from trajectory data)

1. Generate N trajectories $\{x_t\}_{t \in \mathcal{T}}$, where $x_t \in \mathbb{R}^d$ and $\mathcal{T} \subset [0, \tau]$ has cardinality T .
2. Select the time diffusion strength a according to the heuristic in section 4.1.
3. Construct the $N \times N$ matrices M^t, D^t for $t \in \mathcal{T}$ as in section 6.
4. For the current choice of a construct \mathbf{M} and \mathbf{D} as in section 6, and solve the inflated dynamic Laplacian eigenproblem $\mathbf{D}\mathbf{w} = \Lambda \mathbf{M}\mathbf{w}$.
5. Classify eigenfunctions as spatial or temporal by computing the temporal variance of spatial means, as in section 4.2.
6. If all leading nontrivial eigenfunctions are temporal, increase a and return to step 4. Aim for a value of a with a small number of temporal eigenvalues early in the spectrum with most eigenvalues being spatial.
7. Plot the slicewise norms $\|F_k^{\text{spat}}(t, \cdot)\|_2$ vs time t as in Figure 5; large values indicate periods of coherence for the features encoded in F_k^{spat} , while zero or near-zero values indicate strong mixing of the features encoded in F_k^{spat} .
8. Apply SEBA to a collection of leading spatial eigenfunctions F_k^{spat} , $k = 1, \dots, K$, where K is determined by a spectral gap or other means, to produce a family of SEBA functions S_k , $k = 1, \dots, K$.
9. The eigenfunction families $F_k^{\text{spat}}(t, \cdot)$ or SEBA function families $S_k(t, \cdot)$ may be spatially plotted in the pullback space \mathbb{M}_0 or the co-evolved space \mathbb{M}_1 as in Figure 9.

Acknowledgments

We thank Kathrin Padberg-Gehle for initial discussions on the idea of relaxing materiality, Ben Goldys and Georg Gottwald for helpful discussions on SDEs on manifolds, Carsten Gräser and Martin Stahn for discussions on the H^2 -regularity in Appendix B.1, and Shane Keating for pointing us to the Childress–Soward (“cat’s eye”) flow. We further thank Marcus Weber and Konstantin Fackelday for discussions on forward and backward transfer operators.

The research of GF was partially supported by an Australian Research Council Discovery Project. PK thanks the UNSW School of Mathematics and Statistics for support through the Distinguished Visitor Scheme, where much of this research was carried out. PK was partially supported by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 “Scaling Cascades in Complex Systems”, Project Number 235221301, Project A01 “Coupling a multiscale stochastic precipitation model to large scale atmospheric flow dynamics”. Both authors acknowledge the Australian-German Joint Research Cooperation Scheme, which supported visits to UNSW and FU Berlin.

A. Eigenvalue bounds and asymptotics

Proof of Theorem 7.

Part 1 If $k = 1$, then $\lambda_1^D = \Lambda_{1,a}^{\text{spat}} = 0$ for all $a > 0$. For $k \geq 2$, note that a minimizer of (33) is the eigenfunction f_k . Let $\tilde{F}_i(t, x) := f_i(x)$ for $(t, x) \in \mathbb{M}_0$ and $i \geq 1$. Let $H^1 := H^1(\mathbb{M}_0)$ denote the Sobolev space of weakly differentiable $L^2(\mathbb{M}_0)$ functions with derivatives in $L^2(\mathbb{M}_0)$ and set $H_{\text{spat}}^1 = H^1(\mathbb{M}_0) \cap \mathbb{S}_0^{\text{spat}}$. Note that $\tilde{F}_i \in H_{\text{spat}}^1$ with $\partial_t \tilde{F}_i \equiv 0$. We define $\mathbb{S}'_k := \text{span}\{\tilde{F}_1, \dots, \tilde{F}_k\} \subset H_{\text{spat}}^1$, and note that for simplicity we will denote $\nabla_{g_t} F(t, \cdot)$ by $\nabla_{g_t} F$. Then, by the (Courant–Fischer) min-max theorem for self-adjoint operators with no essential spectrum [Tes09, sec. 4.3] applied to the spatial modes we obtain

$$\begin{aligned} -\Lambda_{k,a}^{\text{spat}} &= \min_{\substack{\mathbb{S} \subset H_{\text{spat}}^1 \\ \dim \mathbb{S} = k}} \max_{\substack{F \in \mathbb{S} \\ \|F\|_{L^2} = 1}} \iint_{\mathbb{M}_0} a^2 (\partial_t F)^2 + \|\nabla_{g_t} F\|_{g_t}^2 \, d\ell \, dt \\ &\leq \max_{\substack{F \in \mathbb{S}'_k \\ \|F\|_{L^2} = 1}} \iint_{\mathbb{M}_0} \underbrace{a^2 (\partial_t F)^2}_{=0 \text{ for } F \in \mathbb{S}'_k} + \|\nabla_{g_t} F\|_{g_t}^2 \, d\ell \, dt \\ &= \max_{\substack{F \in \mathbb{S}'_k \\ \|F\|_{L^2} = 1}} \iint_{\mathbb{M}_0} \|\nabla_{g_t} F\|_{g_t}^2 \, d\ell \, dt \stackrel{(33)}{=} -\lambda_k^D, \end{aligned}$$

where the inequality on the second line follows from bounding the minimum by the particular subspace \mathbb{S}'_k , and the last equality comes from noting that the maximizer of the expression on its left-hand side is \tilde{F}_k .

Part 2 The result for $\Lambda_{k,a}^{\text{temp}}$ is obvious using the explicit formula for $\Lambda_{k,a}^{\text{temp}}$. For $\Lambda_{k,a}$, the min-max characterisation states that

$$-\Lambda_{k,a} = \min_{\substack{\mathbb{S} \subset H^1 \\ \dim \mathbb{S} = k}} \max_{\substack{F \in \mathbb{S} \\ \|F\|_{L^2} = 1}} \iint_{\mathbb{M}_0} a^2 (\partial_t F)^2 + \|\nabla_{g_t} F\|_{g_t}^2 \, d\ell \, dt. \quad (57)$$

Because the integrand is nondecreasing in a , the result follows. The argument for $\Lambda_{k,a}^{\text{spat}}$ is similar.

Part 3 Obvious, using the explicit formula for $\Lambda_{k,a}^{\text{temp}}$.

Part 4 For $a > 0$, let $F_a \in \mathbb{S}_k^{\text{spat}}$ be the minimiser of (36) of unit norm in $L^2(\mathbb{M}_0)$. By parts 1 and 2, we know that $a \mapsto \Lambda_{k,a}^{\text{spat}}$ is a nonincreasing function bounded below by λ_k^D . By (38), and (36), we must therefore have that

$$\int_{\mathbb{M}_0} (\partial_t F_a)^2 \rightarrow 0 \quad \text{as } a \rightarrow \infty. \quad (58)$$

Since H^1 is reflexive and $\sup_{a \geq 1} \|F_a\|_{H^1} < \infty$, by Banach–Alaoglu there is a weak accumulation point $F_* \in H^1$ along a sequence $a_i \uparrow \infty$:

$$F_{a_i} \rightharpoonup F_* \in H^1 \text{ as } a_i \rightarrow \infty. \quad (59)$$

We show that $\partial_t F_* \equiv 0$ in the weak sense. For all $\phi \in C^\infty(\mathbb{M}_0)$ we have that

$$\left| \int_{\mathbb{M}_0} \partial_t F_* \phi \right| \stackrel{(59)}{=} \lim_{i \rightarrow \infty} \left| \int_{\mathbb{M}_0} \partial_t F_{a_i} \phi \right| \leq \lim_{i \rightarrow \infty} \|\partial_t F_{a_i}\|_{L^2} \|\phi\|_{L^2} \stackrel{(58)}{=} 0.$$

Denote by $\Lambda_{k,*}$ the value obtained by inserting F_* into (36). Let us denote by $(\cdot, \cdot)_{H^1}$ the L^2 scalar product of the first weak derivatives. We now have that

$$\begin{aligned} -\Lambda_{k,*} &= \iint_{\mathbb{M}_0} \|\nabla_{G_{0,a}} F_*\|_{G_{0,a}}^2 d\ell dt \\ &= (F_*, F_*)_{H^1} && \text{by } \partial_t F_* = 0 \\ &= \lim_i (F_*, F_{a_i})_{H^1} && \text{by (59)} \\ &\leq \lim_i (F_*, F_*)_{H^1}^{1/2} (F_{a_i}, F_{a_i})_{H^1}^{1/2} && \text{by } L^2 \text{ Cauchy–Schwarz} \\ &= (-\Lambda_{k,*})^{1/2} \lim_i \left(\iint_{\mathbb{M}_0} (\partial_t F_{a_i})^2 + \|\nabla_{g_t} F_{a_i}\|_{g_t}^2 d\ell dt \right)^{1/2} \\ &\leq (-\Lambda_{k,*})^{1/2} \lim_i \left(\iint_{\mathbb{M}_0} a_i^2 (\partial_t F_{a_i})^2 + \|\nabla_{g_t} F_{a_i}\|_{g_t}^2 d\ell dt \right)^{1/2} \text{ for } a_i \geq 1 \\ &= (-\Lambda_{k,*})^{1/2} \lim_i (-\Lambda_{k,a_i})^{1/2}. \end{aligned}$$

By part 2, $a \mapsto \Lambda_{k,a}$ is monotone, and so we obtain for the entire sequence

$$\Lambda_{k,*} \geq \lim_{a \rightarrow \infty} \Lambda_{k,a} \geq \lim_{a \rightarrow \infty} \Lambda_{k,a}^{\text{spat}} \stackrel{\text{part 1}}{\geq} \lambda_k^D. \quad (60)$$

Let us consider the previous constructions for all eigenvalues $\Lambda_{j,a}$ for $j = 1, \dots, k$. For the associated eigenfunctions, we obtain as in (59) the weak limits $F_{j,*}$, that are constant in the temporal coordinates. We define $\mathbb{S}'_k := \text{span}\{F_{1,*}, \dots, F_{k,*}\}$, $\mathbb{S}^D := \{F(t, x) = f(x) : f \in H^1(M)\}$ and obtain in complete analogy to the computation in part 1 that

$$\begin{aligned} -\Lambda_{k,*} &= \iint_{\mathbb{M}_0} \|\nabla_{G_{0,a}} F_{k,*}\|_{G_{0,a}}^2 d\ell dt \\ &= \iint_{\mathbb{M}_0} \|\nabla_{g_t} F_{k,*}\|_{g_t}^2 d\ell dt && \text{since } \partial_t F_{k,*} = 0 \\ &= \max_{\substack{F \in \mathbb{S}'_k \\ \|F\|_{L^2} = 1}} \iint_{\mathbb{M}_0} \|\nabla_{g_t} F\|_{g_t}^2 d\ell dt \\ &\geq \min_{\substack{\mathbb{S} \subset \mathbb{S}^D \\ \dim \mathbb{S} = k}} \max_{\substack{F \in \mathbb{S} \\ \|F\|_{L^2} = 1}} \iint_{\mathbb{M}_0} \|\nabla_{g_t} F\|_{g_t}^2 d\ell dt = -\lambda_k^D, \end{aligned}$$

where the inequality follows from $\mathbb{S}'_k \subset \mathbb{S}^D$. With (60), the claim follows. \square

B. Properties of the surrogate model

B.1. Solution of the surrogate problem

Proof of Proposition 8. By classical theory⁶ we have for the solution of the eigenvalue equation that $u \in H_{\text{loc}}^2(0,1)$. Further, by the Sobolev embedding theorem [Eva10, Thm. 5.6.6] in one dimension (implying that H^2 is continuously embedded in C^1) we obtain that u is continuously differentiable on $(0,1)$.

With the ansatz $u(t) = \alpha e^{\omega t}$ in (42) we obtain

$$\omega^2 = \begin{cases} \omega_z^2 = 2\frac{\nu+z}{a^2}, & t \in [0, p], \\ \omega_Z^2 = 2\frac{\nu+Z}{a^2}, & t \in [p, 1], \end{cases}$$

thus the general solution has the form

$$u(t) = \begin{cases} u_z(t) := \alpha_1 e^{\omega_z t} + \alpha_2 e^{-\omega_z t}, & t \in [0, p], \\ u_Z(t) := \alpha_3 e^{\omega_Z t} + \alpha_4 e^{-\omega_Z t}, & t \in [p, 1]. \end{cases}$$

The Neumann boundary conditions translate into

$$t = 0 : \alpha_1 \omega_z - \alpha_2 \omega_z = 0, \quad t = 1 : \alpha_3 \omega_Z e^{\omega_Z} - \alpha_4 \omega_Z e^{-\omega_Z} = 0,$$

that is $\alpha_1 = \alpha_2$ and $\alpha_4 = \alpha_3 e^{2\omega_Z}$. Now we require continuity at the interface $t = p$, i.e.,

$$\alpha_1 (e^{\omega_z p} + e^{-\omega_z p}) = \alpha_3 (e^{\omega_Z p} + e^{2\omega_Z} e^{-\omega_Z p}),$$

equivalently $\alpha_1 \cosh(\omega_z p) = \alpha_3 e^{\omega_Z} \cosh(\omega_Z(1-p))$. Solving this for α_3 and substituting into $u(t)$ (replacing the free constant α_1 by a general constant $\alpha/2$) gives $u_z(t) = \alpha \cosh(\omega_z t)$ and

$$u_Z(t) = \frac{\alpha}{2} \frac{\cosh(\omega_z p)}{\cosh(\omega_Z(1-p))} e^{-\omega_Z} (e^{\omega_z t} + e^{2\omega_Z} e^{-\omega_z t}) = \alpha \frac{\cosh(\omega_z p)}{\cosh(\omega_Z(1-p))} \cosh(\omega_Z(1-t)).$$

To obtain the equation characterising the eigenvalues, we invoke the continuity of the derivative of u at $t = p \in (0,1)$, i.e., $u'_z(p) = u'_Z(p)$, and obtain

$$\omega_z \sinh(\omega_z p) = -\omega_Z \frac{\cosh(\omega_z p)}{\cosh(\omega_Z(1-p))} \sinh(\omega_Z(1-p)),$$

which gives $\frac{\omega_z}{\omega_Z} \tanh(\omega_z p) = -\tanh(\omega_Z(1-p))$, i.e., (45). This concludes the proof. \square

⁶Since classical, we will only sketch the steps here. Consider the bilinear form $B(\cdot, \cdot)$ associated to the differential operator $Lu(t) = u''(t) - \rho(t)u(t)$ with homogeneous Neumann boundary conditions. By the Lax–Milgram theorem [Eva10, Thm. 6.2.1] and Poincaré’s inequality [Eva10, Thm. 5.8.1] there is a unique, well-defined solution operator $S : L^2 \rightarrow H^1 \cap \mathbf{1}^\perp$, $f \mapsto u$, of $B(u, v) = \langle f, v \rangle \forall v \in H^1 (*)$. By the Sobolev embedding theorem [Eva10, Thm. 5.6.6] S is compact and thus has countable spectrum. If $u \in H^1$ is an eigenfunction for some eigenvalue $\nu \neq 0$, then u solves $(*)$ with $f = \nu u \in L^2$, and by the regularity results for elliptic equations [Eva10, Thm. 6.3.1] we obtain $u \in H_{\text{loc}}^2$.

B.2. Eigenvalue analysis of the surrogate problem

Without loss, we can take $a > 0$. We consider the situation where the mixing rate is given by the function ρ from (44), i.e., we are here in the setting of Proposition 8. In particular, we will analyse the solutions of (45), looking at the cases $\nu > -z$, $-z > \nu > -Z$, and $-Z > \nu$ separately. Our findings are summarized in Proposition 9 below. To avoid overly complicated formulas, we will suppress the dependence of ω_z and ω_Z on ν in the following.

The case $\nu > -z$ If $\nu > -z$, we have $\omega_z, \omega_Z \in \mathbb{R}$. Further, by the structure of (45) we can then assume $\omega_z, \omega_Z > 0$ without loss, and it follows that every term on the left-hand side of (45) is positive, hence the equation can not have a solution. Thus, every solution of (45) satisfies $\nu < -z$, as $\nu = -z$ can be ruled out by similar arguments.

The case $\nu \in (-Z, -z)$ In this case $\omega_z \in i\mathbb{R}$, $\omega_Z \in \mathbb{R}$. As before, the signs can be chosen such that $\omega_z = i|\omega_z|$, $\omega_Z > 0$. The equation (45) reads in this case as

$$f(\nu) := \tanh(\omega_Z(1-p)) - \frac{|\omega_z|}{\omega_Z} \tan(|\omega_z|p) = 0.$$

With $|\omega_z| = \frac{\sqrt{-2(\nu+z)}}{a}$, $\omega_Z = \frac{\sqrt{2(\nu+Z)}}{a}$, we observe that

- $\nu \mapsto \tanh\left(\frac{\sqrt{2(\nu+Z)}}{a}(1-p)\right)$ is increasing;
- $\nu \mapsto \frac{|\omega_z|}{\omega_Z} = \sqrt{\frac{-\nu-z}{\nu+Z}} = \sqrt{\frac{Z-z}{\nu+Z} - 1}$ is decreasing; and
- $\nu \mapsto \tan\left(\frac{\sqrt{-2(\nu+z)}}{a}p\right)$ is decreasing between its singularities.

In summary, between its singularities, the function f is continuous with values increasing from $-\infty$ to ∞ ; see Figure 6(b). The singularities ν_k^* can be characterised by

$$\frac{\sqrt{-2(\nu_k^* + z)}}{a}p = \frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z} \quad \Longleftrightarrow \quad \nu_k^* = -z - \frac{a^2}{2} \left(\frac{2k+1}{2p} \pi \right)^2, \quad k \in \mathbb{Z}, \quad (61)$$

thus there is exactly one eigenvalue in every open interval defined by two adjacent singularities inside $(-Z, -z)$. The rightmost interval is not bounded by a singularity on the right, but by $-z$, as $f(-z) > 0$ guarantees the existence of a zero of f larger than ν_0^* .

What happens if p is so small that the first singularity of the kind as in (61) satisfies $\nu_0^* < -Z$? We note that then $\tan(|\omega_z|p) > 0$ for $\nu \in (-Z, -z)$, since it is a decreasing function in ν (see above) and it admits the value 0 at $\nu = -z$, while having its first singularity that is left of $-z$, i.e. ν_0^* , smaller than $-Z$. For $\nu \rightarrow -Z^+$ we have $|\omega_z|/\omega_Z \rightarrow +\infty$ and $\tanh(\omega_Z(1-p)) \rightarrow 0$, thus $\lim_{\nu \rightarrow -Z^+} f(\nu) = -\infty$. As f is continuous and monotonically increasing on $(-Z, -z)$, we obtain that it has exactly one zero in $(-Z, -z)$ if $\nu_0^* < -Z$.

For $p \rightarrow 0$ and fixed $\nu \in (-Z, -z)$ we have

$$\tanh(\omega_Z(1-p)) \rightarrow \text{const} > 0, \quad \frac{|\omega_z|}{\omega_Z} \tan(|\omega_z|p) \rightarrow 0.$$

Thus, for any such fixed ν one has that $f(\nu) > 0$ if p sufficiently small, implying that

$$\text{the zero of } f \text{ in } (-Z, -z) \text{ converges to } -Z \text{ as } p \rightarrow 0. \quad (62)$$

For $a \rightarrow 0$ we have that the largest zero of f converges agains $-z$, since by the above there is a zero in the interval $(\nu_0^*, -z)$ and $\lim_{a \rightarrow 0} \nu_0^* = -z$ by (61). To summarize, we have shown:

Proposition 9. *The dominant (largest) eigenvalue ν_0 of the surrogate problem $\frac{a^2}{2}u'' - \rho u = \nu u$ on $(0, 1)$ with homogeneous Neumann boundary conditions and $\rho(t) = z\mathbf{1}_{[0,p]}(t) + Z\mathbf{1}_{(p,1]}(t)$ satisfies:*

$$(a) \quad \nu_0 \in (-Z, -z)$$

$$(b) \quad \lim_{p \rightarrow 0} \nu_0 = -Z$$

$$(c) \quad \lim_{a \rightarrow 0} \nu_0 = -z$$

Additionally, by similar arguments to those in Theorem 7, we expect

$$\lim_{a \rightarrow \infty} \nu_0 = - \int_0^1 \rho(t) dt = -pz - (1-p)Z.$$

The case $\nu < -Z$ In this case

$$f(\nu) = \frac{|\omega_z|}{|\omega_Z|} i \tan(|\omega_z|p) + i \tan(|\omega_Z|(1-p)),$$

and its zeros are sandwiched between singularities of the two trigonometric tangent functions. The situation is depicted in Figure 6(b) for $z = 2$, $Z = 40$, $p = 0.25$, $\frac{a^2}{2} = \frac{1}{\pi^2}$. The red crosses indicate the eigenvalues of (42) computed by a finite difference scheme on a uniform grid of 1000 nodes.

References

- [ACKBV20] F. Andrade-Canto, D. Karrasch, and F. Beron-Vera. Genesis, evolution, and apocalypse of loop current rings. *Physics of Fluids*, 32(11):116603, 2020.
- [AFJ] F. Antown, G. Froyland, and O. Junge. Linear response for the dynamic Laplacian and finite-time coherent sets. To appear in *Nonlinearity*, 2021.
- [AP15] M. R. Allshouse and T. Peacock. Lagrangian based methods for coherent structure detection. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(9):097617, 2015.

- [AT12] M. R. Allshouse and J.-L. Thiffeault. Detecting coherent structures using braids. *Physica D: Nonlinear Phenomena*, 241(2):95–105, 2012.
- [BBC⁺14] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin. The structure and dynamics of multilayer networks. *Physics reports*, 544(1):1–122, 2014.
- [BGT20] C. Blachut and C. González-Tokman. A tale of two vortices: How numerical ergodic theory and transfer operators reveal fundamental changes to coherent structures in non-autonomous dynamical systems. *Journal of Computational Dynamics*, 7(2):369–399, 2020.
- [BM12] M. Budišić and I. Mezić. Geometry of the ergodic quotient reveals coherent structures in flows. *Physica D: Nonlinear Phenomena*, 241(15):1255–1269, 2012.
- [Cha84] I. Chavel. *Eigenvalues in Riemannian geometry*. Academic press, 1984.
- [Che70] J. Cheeger. *A lower bound for the smallest eigenvalue of the Laplacian*. Princeton U. Press, Princeton, 1970.
- [CS89] S. Childress and A. Soward. Scalar transport and alpha-effect for a family of cat’s-eye flows. *Journal of Fluid Mechanics*, 205:99–133, 1989.
- [CSSdS07] D. B. Chelton, M. G. Schlax, R. M. Samelson, and R. A. de Szoeke. Global observations of large oceanic eddies. *Geophysical Research Letters*, 34(15), 2007.
- [DDSRC⁺13] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas. Mathematical formulation of multilayer networks. *Physical Review X*, 3(4):041022, 2013.
- [DJM16] A. Denner, O. Junge, and D. Matthes. Computing coherent sets using the Fokker–Planck equation. *Journal of Computational Dynamics*, 3(2):163, 2016.
- [EA21] A. El Aouni. A hybrid identification and tracking of Lagrangian mesoscale eddies. *Physics of Fluids*, 33(3):036604, 2021.
- [Eva10] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2nd edition, 2010.
- [FF60] H. Federer and W. H. Fleming. Normal and integral currents. *Annals of Mathematics*, pages 458–520, 1960.
- [FHR⁺12] G. Froyland, C. Horenkamp, V. Rossi, N. Santitissadeekorn, and A. S. Gupta. Three-dimensional characterization and tracking of an Agulhas Ring. *Ocean Modelling*, 52:69–75, 2012.
- [FJ18] G. Froyland and O. Junge. Robust FEM-based extraction of finite-time coherent sets using scattered, sparse, and incomplete trajectories. *SIAM Journal on Applied Dynamical Systems*, 17(2):1891–1924, 2018.
- [FK15] G. Froyland and E. Kwok. Partitions of networks that are robust to vertex permutation dynamics. *Special Matrices*, 3(1):22–42, 2015.

- [FK17] G. Froyland and P. Koltai. Estimating long-term behavior of periodically driven flows without trajectory integration. *Nonlinearity*, 30(5):1948, 2017.
- [FK20] G. Froyland and E. Kwok. A dynamic Laplacian for identifying Lagrangian coherent structures on weighted Riemannian manifolds. *Journal of Nonlinear Science*, 30:1889—1971, 2020.
- [FKN⁺19] K. Fackeldey, P. Koltai, P. N  vir, H. Rust, A. Schild, and M. Weber. From metastable to coherent sets—time-discretization schemes. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(1):012101, 2019.
- [FKS20] G. Froyland, P. Koltai, and M. Stahn. Computation and optimal perturbation of finite-time coherent sets for aperiodic flows without trajectory integration. *SIAM Journal on Applied Dynamical Systems*, 19(3):1659–1700, 2020.
- [FPG14] G. Froyland and K. Padberg-Gehle. Almost-invariant and finite-time coherent sets: directionality, duration, and diffusion. In *Ergodic Theory, Open Dynamics, and Coherent Structures*, pages 171–216. Springer, 2014.
- [Fro13] G. Froyland. An analytic framework for identifying finite-time coherent sets in time-dependent dynamical systems. *Physica D: Nonlinear Phenomena*, 250:1–19, 2013.
- [Fro15] G. Froyland. Dynamic isoperimetry and the geometry of Lagrangian coherent structures. *Nonlinearity*, 28(10):3587, 2015.
- [FRS19] G. Froyland, C. P. Rock, and K. Sakellariou. Sparse eigenbasis approximation: Multiple feature extraction across spatiotemporal scales with application to coherent set identification. *Communications in Nonlinear Science and Numerical Simulation*, 77:81–107, 2019.
- [FSM10] G. Froyland, N. Santitissadeekorn, and A. Monahan. Transport in time-dependent dynamical systems: Finite-time coherent sets. *Chaos*, 20(4):043116, 2010.
- [Fu06] L.-L. Fu. Pathways of eddies in the south atlantic ocean revealed from satellite altimeter observations. *Geophysical Research Letters*, 33(14), 2006.
- [GD20] D. Giannakis and S. Das. Extraction and prediction of coherent patterns in incompressible flows through space–time Koopman analysis. *Physica D: Nonlinear Phenomena*, 402:132211, 2020.
- [GDGGG⁺13] S. G  mez, A. D  az-Guilera, J. G  mez-Garde  es, C. J. P  rez-Vicente, Y. Moreno, and A. Arenas. Diffusion dynamics on multiplex networks. *Physical Review Letters*, 110(2):028701, 2013.
- [GRSK12] P. Grover, S. D. Ross, M. A. Stremler, and P. Kumar. Topological chaos, braiding and bifurcation of almost-cyclic sets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(4):043135, 2012.
- [HBV13] G. Haller and F. Beron-Vera. Coherent Lagrangian vortices: the black holes of turbulence. *Journal of Fluid Mechanics*, 731, 2013.

- [HP98] G. Haller and A. Poje. Finite time transport in aperiodic flows. *Physica D: Nonlinear Phenomena*, 119(3-4):352–380, 1998.
- [Hsu02] E. P. Hsu. *Stochastic analysis on manifolds*, volume 38. American Mathematical Soc., 2002.
- [JMM04] O. Junge, J. E. Marsden, and I. Mezic. Uncertainty in the dynamics of conservative maps. In *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, volume 2, pages 2225–2230. IEEE, 2004.
- [KAB⁺14] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. Multilayer networks. *Journal of complex networks*, 2(3):203–271, 2014.
- [KK20] D. Karrasch and J. Keller. A geometric heat-flow theory of Lagrangian coherent structures. *Journal of Nonlinear Science*, 30:1849–1888, 2020.
- [Kro91] M. Krol. On the averaging method in nearly time-periodic advection-diffusion problems. *SIAM Journal on Applied Mathematics*, 51(6):1622–1637, 1991.
- [KS21] D. Karrasch and N. Schilling. A Lagrangian perspective on nonautonomous advection-diffusion processes in the low-diffusivity limit. *arXiv preprint arXiv:2102.04777*, 2021.
- [Lab15] O. Lablée. *Spectral Theory in Riemannian Geometry*. European Mathematical Society, 2015.
- [MB14] T. Ma and E. M. Bollt. Differential geometry perspective of shape coherence and curvature evolution by finite-time nonhyperbolic splitting. *SIAM Journal on Applied Dynamical Systems*, 13(3):1106–1136, 2014.
- [MMP84] R. MacKay, J. Meiss, and I. Percival. Transport in Hamiltonian systems. *Physica D: Nonlinear Phenomena*, 13(1-2):55–81, 1984.
- [MOR20] T. MacMillan, N. T. Ouellette, and D. H. Richter. Detection of evolving Lagrangian coherent structures: A multiple object tracking approach. *Physical Review Fluids*, 5(12):124401, 2020.
- [NPGR21] M. Ndour, K. Padberg-Gehle, and M. Rasmussen. Spectral early-warning signals for sudden changes in time-dependent flow patterns. *Fluids*, 6(2):49, 2021.
- [Ott89] J. M. Ottino. *The kinematics of mixing: stretching, chaos, and transport*, volume 3. Cambridge university press, 1989.
- [Pav14] G. A. Pavliotis. *Stochastic processes and applications: diffusion processes, the Fokker–Planck and Langevin equations*, volume 60. Springer, 2014.
- [Pie91] R. T. Pierrehumbert. Chaotic mixing of tracer and vorticity by modulated travelling Rossby waves. *Geophysical & Astrophysical Fluid Dynamics*, 84(1-4):285–319, 1991.
- [PS08] G. Pavliotis and A. Stuart. *Multiscale methods: averaging and homogenization*. Springer Science & Business Media, 2008.

- [PY93] R. T. Pierrehumbert and H. Yang. Global chaotic mixing on isentropic surfaces. *Journal of the atmospheric sciences*, 50(15):2462–2480, 1993.
- [RKLW90] V. Rom-Kedar, A. Leonard, and S. Wiggins. An analytical study of transport, mixing and chaos in an unsteady vortical flow. *Journal of Fluid Mechanics*, 214:347 – 394, 1990.
- [RSPB11] I. I. Rypina, S. Scott, L. J. Pratt, and M. G. Brown. Investigating the connection between complexity of isolated trajectories and Lagrangian coherent structures. *Nonlinear Processes in Geophysics*, 18(6):977–987, 2011.
- [SKJ21] N. Schilling, D. Karrasch, and O. Junge. Heat-content and diffusive leakage from material sets in the low-diffusivity limit. *arXiv preprint arXiv:2102.08311*, 2021.
- [SLM05] S. C. Shadden, F. Lekien, and J. E. Marsden. Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Physica D*, 212(3):271–304, 2005.
- [Tes09] G. Teschl. *Mathematical Methods in Quantum Mechanics: With Applications to Schrödinger Operators*, volume 99 of *Graduate Studies in Mathematics*. American Mathematical Society, 2009.
- [Tes12] G. Teschl. *Ordinary differential equations and dynamical systems*, volume 140. American Mathematical Soc., 2012.
- [Thi03] J.-L. Thiffeault. Advection–diffusion in Lagrangian coordinates. *Physics Letters A*, 309(5-6):415–422, 2003.